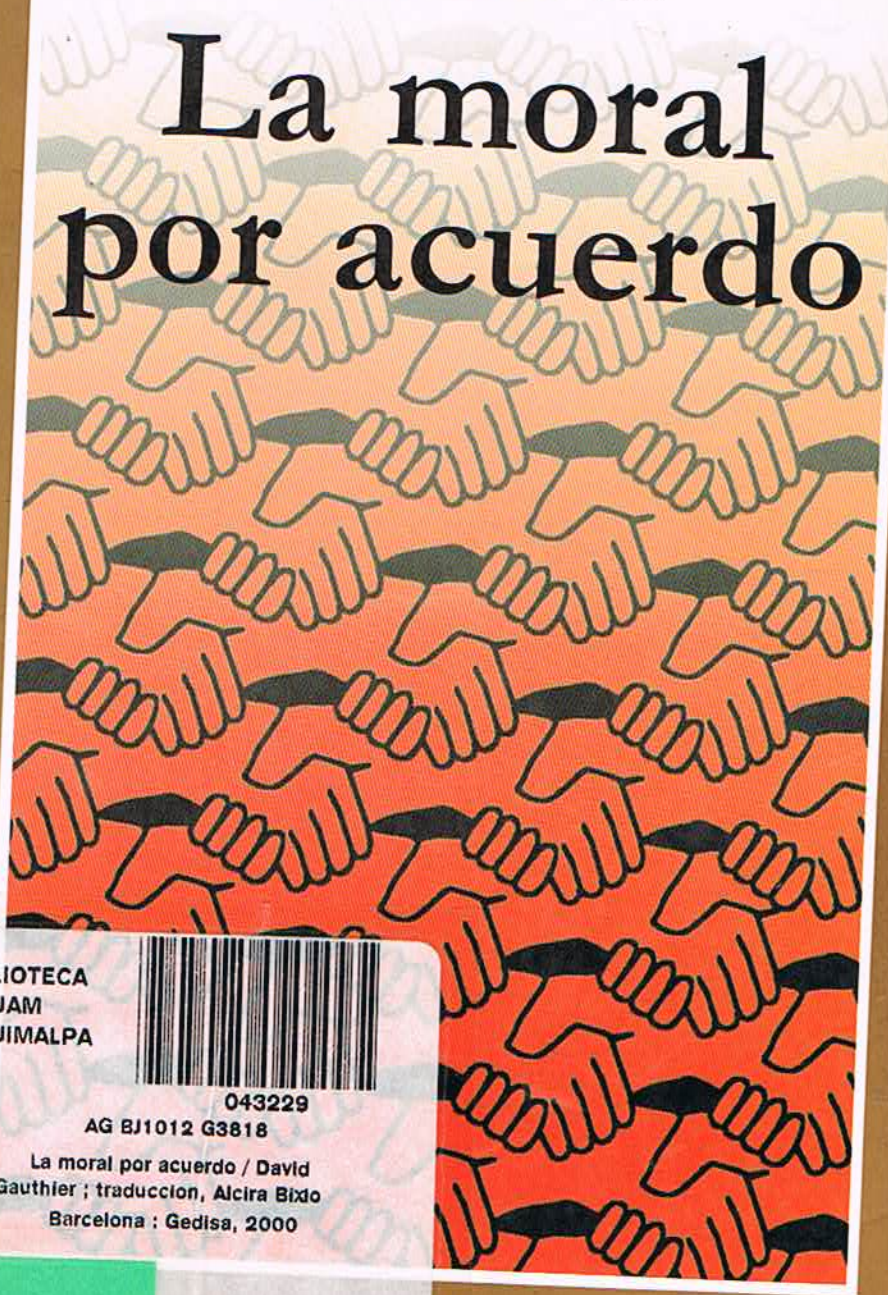
Filosofía del Derecho

David Gauthier

La moral por acuerdo



BIBLIOTECA
UAM
CUAJIMALPA

043229
AG BJ1012 G3818
La moral por acuerdo / David
Gauthier ; traduccion, Alcira Bixio
Barcelona : Gedisa, 2000

gedisa
editorial

Filosofía del Derecho
Serie CLA·DE·MA

gedisa
editorial

David Gauthier
La moral por acuerdo

«... Nuestra teoría debe generar principios racionales necesarios para la elección sin introducir presupuestos morales previos, limitaciones en la búsqueda del interés o el beneficio individual que, siendo imparciales, satisfagan lo que tradicionalmente se entiende por moral. ... Lo que nos interesa es validar la concepción que entiende la moral como un conjunto de restricciones racionales e imparciales que limitan la búsqueda del interés personal sin defender algún código moral particular ...

Tratar de establecer la racionalidad de las restricciones morales no es en sí misma una tarea novedosa ... Quienes se han dedicado a ella apelaron a una concepción de racionalidad práctica, derivada de Kant, completamente diferente de la nuestra (puesto que) ... incluye ya la dimensión moral de imparcialidad que nosotros tratamos de generar.» (David Gauthier)

Gauthier muestra que su planteamiento no sólo asegura el beneficio mutuo y la justicia, con lo cual satisface las normas de la moral, sino además que cada sujeto puede realmente esperar una mayor utilidad. Al resolver lo que parece una paradoja, el autor establece una moral basada en los firmes fundamentos de la razón.

«... El punto más alto alcanzado por el autor y una obra mayor de la moral contemporánea y de la filosofía política ...»

J.E.J. Altham en *The Times Hibger Education Supplement*

«La moral por acuerdo *es quizás la contribución más importante hecha a la ética contractual desde la* Teoría de la justicia *de Rawls.»*

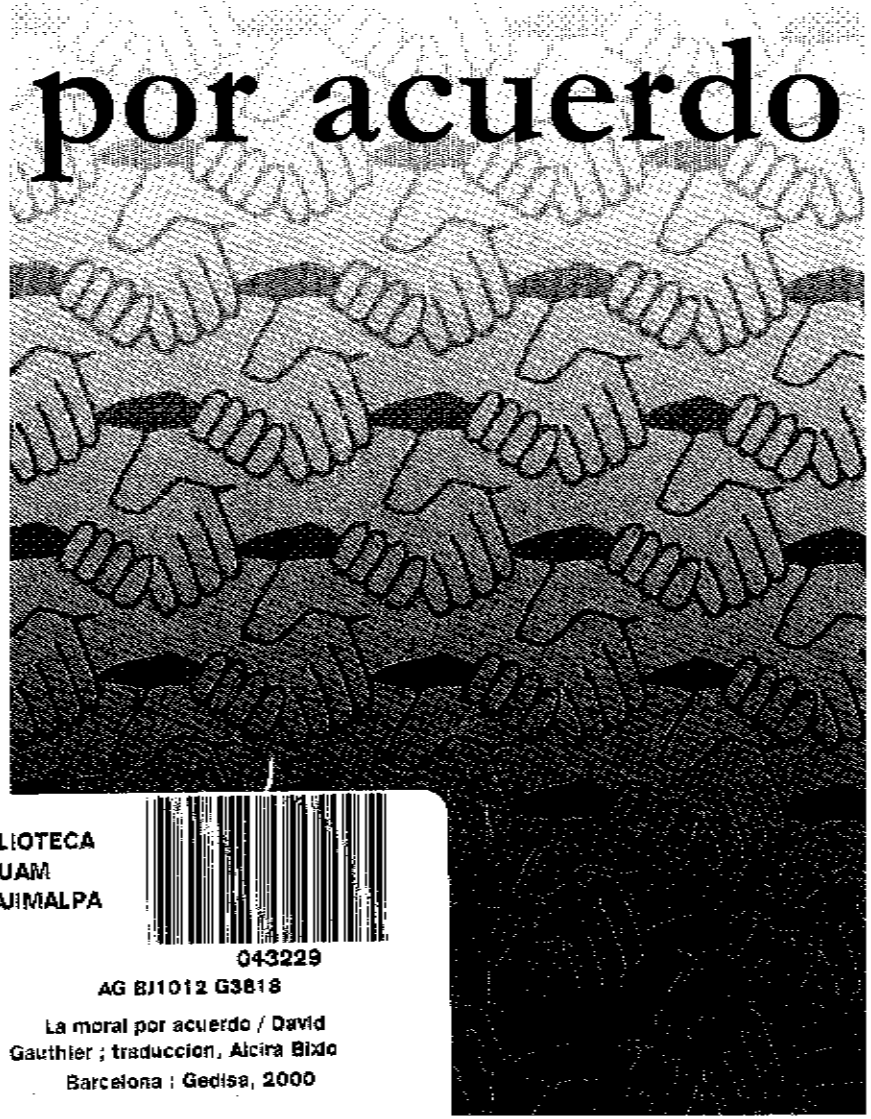
Gregory S. Kavka en *Mind*

David Gauthier es catedrático de filosofía de la Universidad de Pittsburgh, y autor, entre muchos otros textos, de *The Logic of Leviathan.*

ISBN 84-7432-534-X
9 788474 325348
302408

David Gauthier

La moral por acuerdo



BIBLIOTECA
UAM
CUAJIMALPA

043229

AG BJ1012 G3818

La moral por acuerdo / David
Gauthier ; traduccion, Alcira Bixio
Barcelona : Gedisa, 2000

gedisa
editorial

David Gauthier

LA MORAL POR ACUERDO

Serie Cla•De•Ma
Filosofía del derecho

Cla•De•Ma/Derecho

Colección dirigida por
Ernesto Garzón Valdés y Jorge F. Malem Seña

En la actualidad, la filosofía del derecho ya no es un discurso hermético y reservado a los especialistas de cada una de sus parcelas. Además, las modernas herramientas de análisis no sólo abren nuevos accesos metodológicos a los juristas mismos, sino que inciden también en cuestiones que afectan las responsabilidades acerca de nueva civilización, como son las transgresiones de los derechos humanos o la desproporción entre el crecimiento de la riqueza y su necesaria distribución. La colección pretende contribuir al debate que se está desarrollando en la Filosofía del derecho en el ámbito de la lengua castellana con la publicación de obras y recopilaciones de ensayos de autores internacionales cuyos planteamientos merecen una mayor difusión en nuestra área cultural. La intención es ofrecer un panorama lo más amplio posible y dar preferencia a propuestas abiertas a la discusión teórica más que presentar soluciones definitivas.

JORGE MALEM SEÑA	*Globalización, comercio internacional y corrupción*
MICHAEL SANDEL	*El liberalismo y los límites de la justicia*
A. AARNIO, E. GARZÓN VALDÉS, J. UUSITALO *compiladores*	*La normatividad del derecho*
CARLOS SANTIAGO NINO	*La constitución de la democracia deliberativa*
RODOLFO VÁZQUEZ *compilador*	*Derecho y moral*
OWEN FISS	*La ironía de la libertad de expresión*
DENNIS F. THOMPSON	*La ética política y el ejercicio de cargos públicos*
DAVID LYONS	*Aspectos morales de la teoría jurídica*
	Teoría del derecho
J. G. RIDDALL	*Distinguiendo*
RICCARDO GUASTINI	Estudios de teoría y metateoría del derecho
BRUCE ACKERMAN	*La política del diálogo liberal*
DAVID GAUTHIER	*La moral por acuerdo*
NORBERT HOERSTER	*En defensa del positivismo jurídico*
LUCIAN KERN	*La justicia ¿discurso o mercado?*
HANS PETER MÜLLER	*Los nuevos enfoques de la teoría contractual*

LA MORAL POR ACUERDO

por

David Gauthier

gedisa
editorial

3818

Título del original en inglés:
Morals by agreement
Originalmente publicado por Oxford University Press
© David Gauthier,1986

Traducción: Alcira Bixio

Revisión técnica: Samuel Monder

Composición de originales: Laser Factory

Ilustración de cubierta: Juan Santana

Primera reimpresión: septiembre del 2000, Barcelona

Derechos reservados para todas las ediciones en castellano

© Editorial Gedisa, S.A.
Paseo Bonanova, 9 1°-1ª
08022 Barcelona (España)
Tel. 93 253 09 04
Fax 93 253 09 05
Correo electrónico: gedisa@gedisa.com
http://www.gedisa.com

ISBN: 84-7432-534-X
Depósito legal: B. 39828-2000

Impreso por: Carvigraf
Clot, 31 - Ripollet

Impreso en España
Printed in Spain

Queda prohibida la reproducción total o parcial por cualquier medio de impresión, en forma idéntica, extractada o modificada, en castellano o en cualquier otro idioma.

Indice

PREFACIO 9
ADVERTENCIA AL LECTOR 13

I. Visión general de una teoría 15
Notas 38
II. Elección: razón y valor 41
Notas 88
III. Estrategia: razón y equilibrio 91
Notas 119
IV. El mercado: la libertad que procede de la moral 120
Notas 156
V. Cooperación: negociación y justicia 158
Notas 210
VI. El acatamiento: un modo de restringir la maximización 213
Notas 253
VII. La posición negociadora inicial: los derechos y la estipulación 255
Notas 306
VIII. El punto de Arquímedes 308
Notas 349
IX. Personas, pueblos, generaciones 353
Notas 398
X. El anillo de Giges 401
Notas 430
XI. El individuo liberal 432
Notas 463

INDICE TEMÁTICO 465

Prefacio

La presente indagación comenzó una tarde de noviembre en Los Ángeles cuando, mientras buscaba yo a tientas las palabras que me permitieran expresar la peculiar relación que existe entre moral y conveniencia, alguien me presentó el dilema del prisionero. (El lector que no esté familiarizado con él podrá encontrarlo en la sección 3.2 del capítulo III). Casi quince años después reflexiono sobre el curso de un viaje que aún no he completado ni puedo completar, pero que encuentra temporariamente un puerto en este libro.

El dilema del prisionero no resuelve un problema, antes bien lo plantea. Se trata del problema de la racionalidad práctica, entendida en términos de maximización, y según creo, queda resuelto en el capítulo VI.

Este problema resultó ser el segundo de tres problemas medulares que exigen una resolución para poder internarnos en la indagación que desarrollo en este libro. El primero era formular el principio de cooperación racional que, según creo, es esencial para la moral. En mi perspectiva, este principio se relaciona con un acuerdo o una negociación racional, concepto que desarrollo en el capítulo V. El segundo problema consistía en demostrar la racionalidad de aceptar ese principio, que resultó ser el problema de la conducta racional que se plantea en el dilema del prisionero. Y el tercero era determinar la posición inicial apropiada partiendo de la cual puede dar resultados la cooperación; lo cual exige mostrar la racionalidad de aceptar la estipulación que propone Locke para el acuerdo inicial. (El lector que no esté familiarizado con el concepto de la estipulación lockeana puede encontrarlo en la sección 3.1 del capítulo VII). El tercer problema resultó ser el más rebelde; partiendo de la idea inicial de una teoría moral contractual, que atrapó mi imaginación en 1966, tuve que dejar pasar unos trece años para llegar a comprender claramente el papel que desempeña la estipulación.

Durante esos años publiqué varios artículos en los que desarrollé lo que hoy considero protoversiones de partes de la presente teoría. Un lector familiarizado con esos artículos quizá descubra en ellos argumentos y actitudes que el presente análisis contradice o modifica. Prefiero pensar que eso representa un progreso de mi indagación.

Quizá los cambios de actitud merezcan una observación adicional. He tenido, y aun hoy tengo, sentimientos algo confusos en cuanto a una moral basada en el acuerdo. En realidad, alguna vez pensé publicar la mayor parte del presente libro como un estudio de un conjunto de interconexiones conceptuales, sin abrigar la pretensión de que el conjunto constituya la correcta teoría moral. Pero ahora sí lo pretendo. Incorporar la estipulación lockeana y la idea del individuo liberal (expuesta en el último capítulo) fue algo que aligeró mis primeras preocupaciones. Pero quizá lo más importante sea que la concepción de la racionalidad práctica (que admito está en la raíz de mi argumento) me parece la única capaz de soportar el examen crítico, y la teoría moral que desarrollo luego me parece, en líneas generales si no ya en cada detalle, la única compatible con esa concepción de racionalidad. Sin embargo, como podrían recordarme Richard Rorty y Alasdair MacIntyre, quizá carezco del vocabulario necesario para hablar con la suficiente claridad acerca de la moral.

El primer borrador de la mayor parte del libro fue escrito en Aix-en Provence entre 1979 y 1980; agradezco el apoyo financiero que me brindaron durante ese período la Universidad de Toronto y el Consejo de Investigación de Ciencias Sociales y Humanidades de Canadá. Escribí el segundo borrador de los capítulos II al X en Toronto en la primavera de 1982; los capítulos finales se rehusaron a aparecer hasta el verano de 1983. Finalmente escribí el presente borrador en Pittsburgh entre enero y mayo de 1984 y agradezco a la Universidad de Pittsburgh por haberme relevado de mis deberes de profesor durante ese período y a Ruth Durst por haber pasado a máquina el manuscrito.

Si pretendiera agradecer nombrándolas con nombre y apellido a todas las personas que contribuyeron a que esta indagación llegara a ser lo que es, inevitablemente mi lista

pecaría por omisión. Muchos, muchísmos colegas y estudiantes de las universidades de Toronto y Pittsburgh y de otras universidades que visité o cuyos artículos leí, y que traté en reuniones profesionales, me ayudaron a reconcebir mis ideas y a reformular mis pensamientos. A todos ellos, mi gratitud. A Clark Glymour, quien me señaló la conveniencia de evitar los horribles "él/ella" de los primeros borradores recurriendo a un artificio fortuito apropiadamente programado, seguramente puedo agregarle además de mi propio agradecimiento, el de los lectores. También le debo mi gratitud a Howard Sobel, que entre todos aquellos que me brindaron ayuda filosófica ocupa el primer lugar, pues fue él quien me presentó el dilema del prisionero y me introdujo en las ideas básicas de la teoría de los juegos, aunque creo que la gratitud sola no sería suficiente.

Y así llego a un final sabiendo que constituye también un principio, porque seguramente muy pronto me encontraré embarcado nuevamente en la tarea de comprender cómo se relacionan la moral y la racionalidad.

DAVID GAUTHIER
Pittsburgh, 24 de mayo de 1984.

Advertencia al lector

Uno debe comenzar por el comienzo y terminar por el final. Pero *en route*, algunas secciones y subsecciones pueden omitirse en una primera lectura, a fin de captar más fácilmente la línea principal de la argumentación. Además, unas pocas secciones están destinadas únicamente a aquellas personas que tienen un especial interés en las teorías de los juegos y de la elección racional.

Después de obtener una visión de conjunto, que se presenta en el capítulo 1, el lector que desee pasar más rápidamente a las principales tesis de la teoría moral puede omitir las secciones 2 y 3 del capítulo II en las que se trata bastante detalladamente la cuestión de la preferencia, la sección 2 del capítulo III (de interés únicamente para los especialistas, según sospecho) y la sección 4 que da fin al capítulo IV, dedicada a las reacciones utilitaristas y marxistas al mercado. Aunque ciertas referencias posteriores remiten a algunas de estas secciones, en una primera lectura, el hecho de saltarlas no ha de impedir la comprensión general.

Los capítulos V, VI y VII constituyen el centro argumental del libro. Pero aun en ellos es posible hacer algunas omisiones. El lector puede preferir pasar de la sección 1 del capítulo V a la 2.3 y omitir así la discusión referente a las alternativas de la negociación rechazadas y dejar de lado la discusión especializada sobre la alternativa rechazada dentro de la negociación. También pueden omitirse en una primera lectura las secciones 1.2 y 1.3, sobre el análisis de Hobbes y el comentario final sobre el altruismo recíproco de la sección 3.3 del capítulo VI. En el capítulo VII puede ignorarse sin inconvenientes sólo el breve tratamiento del tema de la amenaza de la sección 2.2.

El argumento positivo del capítulo VIII se encuentra en las secciones 1 y 4, las discusiones críticas de Harsanyi y Rawls que aparecen en las secciones intermedias 2 y 3 pueden posponerse

para otra lectura. El capítulo IX esboza las aplicaciones de la teoría y puede omitirse por completo si se prefiere seguir la línea argumental principal; aunque la sección 1 puede suministrar una visión útil de las diferentes aplicaciones y la 5.1, que ofrece los criterios elegidos para clasificar los estilos de vida, es. particularmente importante. En el capítulo X quizás el lector prefiera pasar de la sección 2.1 a la 2.5 y omitir así la discusión detallada del hombre económico. Creo que el capítulo XI debería leerse sin omisiones.

De modo que el material que aconsejo incluir en una primera lectura abarca algo más de los dos tercios de la obra.

I

Visión general de una teoría

1. ¿Qué teoría de la moral puede ser útil a cualquier propósito a menos que pueda demostrar que todos los deberes que recomienda representan también el verdadero interés de cada individuo?[1] David Hume, que fue quien se formuló esta pregunta, parece errado; una teoría semejante sería demasiado útil. Si el deber no es algo superior al interés, la moral sería superflua ¿ Por qué apelar a lo correcto o incorrecto, a lo bueno o lo malo, a la obligación o el deber, si podemos en cambio apelar al deseo o la aversión, al perjuicio o el beneficio, al interés o la ventaja? La apelación a la moral surge al advertir que estas últimas consideraciones no son guías suficientes para hacer lo que debemos hacer. El nada filosófico poeta Ogden Nash captó los supuestos que están en la base de nuestro lenguaje moral de manera más clara que el filósofo Hume, cuando escribió:

> ¡Oh, deber!
> ¿Por qué no tienes el rostro dulce y atractivo?[2]

Podemos lamentar el rostro severo del deber, pero no podemos negarlo. Creemos que sólo por eso algunos pasan por alto el interés o el beneficio que representa una conducta ética.

Pero si bien el lenguaje de la moral puede no ser el del interés, sin duda es el de la razón. De modo que sería mejor preguntarse: ¿Qué teoría de la moral puede servir a cualquier propósito útil, si no puede mostrar que todos los deberes que recomienda se fundan también en la razón de cada individuo? Si los llamamientos de la moral dan lugar a algún efecto práctico y ejercen alguna influencia en nuestra conducta, ello se debe no a que sean un susurro de invitación para nuestros deseos, sino a que persuadan a nuestro intelecto. Supongamos que descubriéramos, como cree el propio Hume, que la razón es impotente en la esfera de la acción, salvo en lo que se refiere a decidir las cuestiones de hecho.[3] O supongamos que descubriéramos que

esa razón sólo es la ayudante de cámara del interés, de modo tal que un llamamiento moral al desechar la conveniencia deba contradecir también la razón. En ambos casos tendríamos que llegar a la conclusión de que la empresa moral, como se la ha concebido tradicionalmente, es imposible.

Decir que nuestro lenguaje moral mantiene una conexión con la razón no equivale a argumentar en favor de la racionalidad de nuestras opiniones morales o de cualquier otra opinión alternativa. El lenguaje moral puede basarse en un supuesto falso.[4] Si los deberes morales se basan en la racionalidad, los emotivistas, que suponen que los llamamientos morales son sólo persuasivos, y los egoístas, que suponen que los llamamientos racionales están limitados por el propio interés, están equivocados.[5] Pero, los deberes morales, ¿están realmente basados en la racionalidad? Eso es lo que trataremos de probar, mostrando que la razón desempeña un papel práctico que está relacionado con el interés individual y que además lo trasciende, de modo tal que los principios morales que prescriben deberes pasando por alto las ventajas o beneficios pueden justificarse racionalmente. Defenderemos la concepción tradicional de la moral como una limitación en la busca del interés personal.

Ahora bien, el error en el que incurre Hume al insistir en que los deberes morales deben representar el verdadero interés de cada individuo, oculta una idea fundamental. La razón práctica está vinculada con el interés o, como hemos llegado a decir, con la utilidad individual, y las restricciones racionales que se anteponen a la busca del interés tienen su base en ese interés que limitan. El deber deja de lado el beneficio, pero la aceptación del deber es verdaderamente beneficiosa. Veremos que esta aparente paradoja encaja en la verdadera estructura de la interacción. Cuando lleguemos a comprender esa estructura, reconoceremos la necesidad de restringir la búsqueda que cada persona hace de su propia utilidad y examinaremos las implicaciones que eso tiene tanto para nuestros principios de acción como para nuestra concepción de la racionalidad práctica. Nuestra indagación nos conducirá a la base racional de una moral, no de normas absolutas sino de limitaciones basadas en el acuerdo.

2.1 Desarrollaremos una teoría de la moral. Nuestra intención es suministrar a la conducta y a los principios morales un marco de justificación y no de explicación. Es decir que desarrollaremos una teoría normativa. Una filosofía completa de la moral necesitaría explicar, y quizá defender, la idea de una teoría normativa. Pero nosotros no haremos eso. En cambio ejemplificaremos la teoría normativa esbozando la teoría de la elección racional. En realidad haremos más que eso. Desarrollaremos una teoría de la moral como parte de la teoría de la elección racional. Sostendremos que los principios racionales necesarios para hacer una elección o para tomar decisiones dentro de las acciones posibles incluyen algo que restringe de manera imparcial la acción del actor que persigue su propio interés. Y eso es lo que identificamos como principios morales.

El estudio de la elección comienza estipulando una concepción clara del valor y la racionalidad en una forma aplicable a las situaciones de elección.[6] La teoría analiza luego la estructura de tales situaciones, de modo que, para cada tipo de estructura identificada, pueda elaborarse la concepción de racionalidad dentro de un conjunto de determinadas condiciones en la que se hace una elección entre las posibles acciones. Estas condiciones se expresan pues como principios precisos de conducta racional, que sirven tanto para la prescripción como para la evaluación crítica. En consecuencia, los principios desempeñan también un papel explicativo en la medida en que las personas en realidad actúan racionalmente.

La parte más simple, más familiar e históricamente primaria de este estudio constituye la médula de la teoría económica clásica y neoclásica, que examina la conducta racional en aquellas situaciones en las cuales el actor conoce con certeza el resultado de cada una de sus posibles acciones. El economista, por supuesto, se ofrece a explicar la conducta y el principal interés de su teoría depende de que tenga aplicaciones explicativas, pero las explicaciones del economista utilizan un modelo de interacción ideal que incluye la racionalidad de los actores entre algunos de sus supuestos. Por lo tanto, la explicación económica se sitúa dentro de un contexto normativo. Y la economía, en su tarea de formular y evaluar las alternativas políticas, no debería dejarnos ninguna duda respecto del carácter profundamente prescriptivo y crítico de la ciencia.

El economista formula una concepción simple y maximizadora de la racionalidad práctica que nosotros analizaremos en el capítulo 2. Pero el supuesto de que el resultado de toda posible elección puede ser conocido con certeza limita seriamente el objetivo del análisis económico y la posibilidad de aplicar su estimación de la razón. La teoría de la decisión bayesiana se aparta de ese supuesto y examina situaciones en las cuales las elecciones implican riesgo o incertidumbre. El teórico de la decisión se ve llevado a extender la estimación de la razón que hace el economista sin dejar de conservar por ello su identificación fundamental de la racionalidad con la maximización.

El análisis de la interacción que hacen tanto la teoría económica como la teoría de la decisión es limitado puesto que ambas consideran únicamente los resultados en relación con las elecciones de un único actor, y consideran las elecciones de los demás como aspectos de las circunstancias de ese actor. La teoría de los juegos supera esa limitación, pues analiza los resultados en relación con un conjunto de elecciones que corresponden a cada una de las personas implicadas en la obtención del resultado. La teoría de los juegos considera las elecciones de un actor que decide basándose en las expectativas de las elecciones que han de hacer los demás, que a su vez deciden basándose en las expectativas de lo que ha de decidir aquél. Puesto que las situaciones que incluyen a un actor único pueden considerarse casos limitados de interacción, la teoría de los juegos apunta a considerar la conducta racional en toda su amplitud. No resulta sorprendente que los logros se relacionen inversamente con los objetivos. La teoría económica como estudio de la conducta racional en situaciones de certeza está esencialmente completa, mientras que la teoría de los juegos está todavía en desarrollo. La teoría de la elección racional es una empresa en marcha que está extendiendo una comprensión básica del valor y la racionalidad a la formulación de principios de conducta racional en una variedad aun más amplia de situaciones.

2.2 La elección racional suministra un modelo de teoría normativa. Uno podría suponer que la teoría moral y la teoría de la elección sólo están relacionadas por estructuras semejantes. Pero, como ya dijimos, desarrollaremos aquí una teoría

moral como parte de la teoría de la elección. Aquellos lectores que conocen los más recientes trabajos de la filosofía moral pueden estar familiarizados con esta empresa; John Rawls ha sostenido que la teoría de la justicia es "quizá la parte más significativa de la teoría de la elección racional", y John Harsanyi trata explícitamente la ética como parte de la teoría de la conducta racional.[7] Pero estas pretensiones son más ambiciosas de lo que pueden asegurar sus resultados. Ni Rawls ni Harsanyi desarrollan la conexión profunda que existe entre la moral y la elección racional que nosotros defenderemos. Una breve comparación ha de dar mayor claridad a nuestro empeño.

Nuestra posición sostiene que en ciertas situaciones que incluyen la interacción con los demás, un individuo elige racionalmente sólo si limita la búsqueda de su propio interés o conveniencia a fin de acatar los principios que expresan la imparcialidad característica de la moral. Para elegir racionalmente uno debe elegir moralmente. Esta es una ambiciosa pretensión. La moral, según mostraremos más adelante, puede aparecer como una coacción racional surgida de premisas no morales de elección racional. Ni Rawls ni Harsanyi sostienen semejante idea. Ni Rawls ni Harsanyi consideran los principios morales como un subconjunto de principios racionales necesarios para la elección.

Rawls sostiene que los principios de justicia son los objetos de una elección racional: la elección que cualquier persona haría si se le pidiera que seleccionara los principios básicos de su sociedad manteniendo ante sus ojos un "velo de ignorancia" que le ocultara todo conocimiento de su propia identidad.[8] Los principios así elegidos no están directamente relacionados con la toma de decisiones individuales.[9] En consecuencia, la aceptación de esos principios debe tener implicaciones en la conducta individual, pero Rawls nunca declara que esas implicaciones incluyen limitaciones racionales en las elecciones individuales. Y tales restricciones pueden ser, empleando la terminología del propio Rawls, limitaciones razonables, pero lo razonable es en sí mismo materia moralmente sustantiva que está más allá de las fronteras de la elección racional.[10]

La idea de Rawls de que los principios de justicia son los objetos de una elección racional en realidad forma parte de

nuestra propia teoría, aunque nosotros representaremos la elección como una negociación o un acuerdo entre personas que necesitan tener conciencia de sus propias identidades. Pero este paralelo entre nuestra teoría y la de Rawls no debe oscurecer la diferencia básica; nuestra pretensión es concebir la moral como un conjunto de principios racionales necesarios para hacer una elección. Nuestro compromiso es mostrar por qué un individuo, al razonar partiendo de premisas no morales, estaría dispuesto a aceptar en sus elecciones las restricciones de la moral.

Aparentemente la teoría de Harsanyi sólo difiere de la de Rawls en cuanto a los principios que una persona elegiría desde detrás del velo de ignorancia. Rawls supone que las personas elegirían los dos bien conocidos principios de la justicia, mientras que Harsanyi supone que las personas elegirían principios basados en general en el utilitarismo.[11] Pero el argumento de Harsanyi en algunos aspectos se acerca más al nuestro; a él le preocupan los principios necesarios para hacer una elección moral y la forma racional de llegar a esos principios. Con todo, los principios de Harsanyi son estrictamente hipotéticos, rigen la elección racional partiendo de una posición imparcial o suponiendo preferencias imparciales, con lo cual sólo son principios válidos para alguien que quiere elegir moral o imparcialmente.[12] Pero Harsanyi no pretende, como nosotros, que haya situaciones en las que un individuo deba elegir moralmente para elegir racionalmente. Para Harsanyi existe una forma racional de elegir moralmente, pero no hay una exigencia racional para hacerlo. Y ésta es otra de las diferencias básicas entre su teoría y la nuestra.

Si ponemos ahora de un lado los puntos de vista de Rawls y Harsanyi —puntos de vista a los que volveremos a remitirnos en los capítulos posteriores—, podemos resumir la significación de las diferencias que hemos esbozado ya. Nuestra teoría debe generar, estrictamente como principios racionales necesarios para la elección y sin introducir supuestos morales previos, limitaciones en la búsqueda del interés o el beneficio individual, limitaciones que, siendo imparciales, satisfagan lo que tradicionalmente se entiende por moral. No suponemos de antemano que tales restricciones o limitaciones imparciales y racionales deban existir. Ni siquiera suponemos que deban existir restricciones racionales, sean o no imparciales. Pretendemos demostrar que existen restricciones racionales y que esas restricciones son imparciales. Luego identificamos la moral con esas restricciones demostradas, pero si su contenido corresponde o no al de los principios morales convencionales, es otra cuestión que no examinaremos en detalle. Sin duda, ha de haber diferencias, quizá significativas, entre las limitaciones imparciales y racionales sostenidas por nuestra argumentación y la moral aprendida de los padres y los pares, de los sacerdotes y los profesores. Pero lo que nos interesa es validar la concepción que entiende la moral como un conjunto de restricciones racionales e imparciales que limitan la busca del interés personal y no defender ningún código moral particular. Además nos importa hacerlo, insistimos, sin incorporar en las premisas de nuestra argumentación ninguna de las concepciones morales que surjan de nuestras conclusiones.

2.3 Tratar de establecer la racionalidad de las restricciones morales no es en sí misma una tarea novedosa, y sus antecedentes son más venerables que el empeño por desarrollar una teoría moral como parte de la teoría de la elección racional. Pero quienes se han dedicado a ello apelaron típicamente a una concepción de racionalidad práctica, derivada de Kant, completamente diferente de la nuestra.[13] En efecto, el modo de entender la razón de esos pensadores incluye ya la dimensión moral de imparcialidad que nosotros tratamos de generar.

Supongamos que se admita que existe una conexión entre la razón y el interés —o la ventaja, el beneficio, la preferencia, la satisfacción o la utilidad individuales, puesto que las diferencias entre estos términos, aunque importantes en otros contextos, no afectan el presente análisis—. Y hasta convengamos en que mientras no se afecten los intereses de los demás, una persona actúa racionalmente si y sólo si busca su máximo interés o beneficio. Habrá quienes nieguen esto, pero lo que nos interesa aquí es aislar la diferencia esencial que existe entre las concepciones opuestas de la racionalidad práctica. Y esa diferencia aparece cuando consideramos la acción racional en la que están implicados los intereses de los demás. Los defensores de la concepción *maximizadora* de la racionalidad, que nosotros respaldamos, sostienen que esencialmente no cambia nada; la

persona racional continúa buscando la máxima satisfacción de sus propios intereses. Por otro lado, quienes defienden lo que llamaremos la concepción *universalista* de la racionalidad insisten en afirmar que lo que le da un carácter racional a la satisfacción de un interés no depende del interés de quien esté en juego. Por lo tanto, la persona racional busca satisfacer todos los intereses. En el presente análisis carece de importancia que se trate de una persona utilitarista que busca la mayor felicidad de la mayor cantidad de gente o una persona que le dedique una consideración independiente a la distribución justa del beneficio entre las personas.

A fin de evitar posibles malentendidos, adviértase que tampoco la concepción de racionalidad exige que las razones prácticas atiendan al propio interés. En la concepción maximizadora no son los intereses en el sí mismo —que consideran a cualquier individuo como objeto—, sino los intereses del sí mismo —sostenidos por uno mismo como sujeto— los que suministran la base para una elección y una acción racionales. En la concepción universalista no son los intereses en cualquier persona —que consideran a cualquier persona como objeto—, sino los intereses de cualquier persona —sostenidos por alguna persona como sujeto— los que proporcionan la base para una elección y una acción racionales. Si yo tengo un interés directo en el bienestar de una determinada persona, en cualquiera de las dos concepciones tengo razones para promover el bienestar de esa persona. Pero el interés de esa persona en su propio bienestar, sólo me proporciona razones semejantes en la concepción universalista.

Ya hemos dicho que tradicionalmente se entiende que la moral implica una restricción imparcial en la busca del interés individual. La justificación de tal limitación no les plantea ningún problema a los defensores de la racionalidad universalista. La exigencia racional de que deben satisfacerse todos los intereses en la máxima extensión posible limita directamente a cada persona en la busqueda de sus propios intereses. La formulación precisa de la restricción ha de depender, por supuesto, de la forma en que se satisfagan los intereses, pero su racionalidad básica es suficientemente clara.

Quienes proponen la concepción universalista de la razón

práctica cumplen por lo tanto fácilmente con la principal tarea de nuestra teoría moral: generar restricciones morales concebidas como racionales. Para ellos la relación que existe entre la razón y la moral queda claramente establecida. En cambio, la tarea que tienen por delante es defender su concepción de racionalidad, puesto que las concepciones maximizadora y universalista no se basan en fundamentos iguales. Entre esas concepciones, la maximizadora posee la cualidad de la debilidad. Cualquier consideración que proporcione una razón para obrar según la concepción maximizadora, proporciona también una razón semejante para la concepción universalista. Pero no puede decirse lo mismo en sentido inverso. Según la concepción universalista todas las personas cuentan, en efecto, con la misma base para hacer una elección racional —los intereses de todos— y ese supuesto de imparcialidad o impersonalidad de la razón, exige una justificación.

Además, y quizás esto sea aun de mayor importancia, la concepción maximizadora de la racionalidad está aceptada casi universalmente y es la que se emplea principalmente en las ciencias sociales.[14] Como ya lo hemos señalado, ésa es la concepción que está en la base de la teoría económica y que se generaliza en la teoría de la decisión y de los juegos. La menor importancia que se le da a esta concepción en las teorías política, sociológica y psicológica, refleja más el poco interés que muestran los estudiosos de esas disciplinas por la racionalidad que la adhesión a una concepción alternativa. No hay duda de que los estudiosos de las ciencias sociales pueden estar equivocados, pero asignamos el peso de la prueba a aquellos que defienden la concepción universalista.

Por consiguiente, al desarrollar una teoría moral dentro de la elección racional, nos inclinamos por la más débil y más ampliamente aceptada de las dos concepciones de la racionalidad que hemos señalado. Evidentemente no debemos suponer que los principios morales que generemos sean idénticos a aquellos que pueden derivar de la concepción universalista. Sus defensores pueden sostener que el modo que ellos tienen de entender la conexión entre la razón y la moral es el correcto, aun cuando acepten que una forma de moral puede basarse en la racionalidad maximizadora. Pero nosotros podemos sugerir, sin defender

aquí nuestra sugerencia, que pocas personas se adherirían a la concepción universalista de la razón práctica si no creyeran que es necesaria para justificar cualquier forma de moral racional. Por lo tanto, la refutación más efectiva de la posición universalista puede ser no tratar de socavar sus elaborados e ingeniosos argumentos, sino elaborar una concepción alternativa de una moral racional basada en los supuestos más débiles de la teoría de la elección racional.

3.1 Una moral basada en el acuerdo parte de un supuesto inicial contra la moralidad entendida como una restricción que cada persona tiene en la búsqueda de su propio interés. Desde ese punto de vista, se concibe a cada persona como un centro independiente de actividad que se esfuerza en dirigir sus facultades y recursos al objetivo de satisfacer sus intereses. Esa persona considera lo que puede hacer, pero inicialmente no establece ninguna distinción entre lo que le está permitido hacer y lo que no lo está. ¿Cómo puede entonces esa persona llegar a reconocer esa distinción? ¿Cómo puede una persona llegar a reconocer una dimensión moral ante una elección, si lo moral no está inicialmente presente?

Una moral basada en el acuerdo ofrece una racionalidad que parte de un contrato establecido a fin de distinguir qué le está permitido hacer y qué no a cada persona. Los principios morales se presentan como los objetos de un acuerdo *ex ante* completamente voluntario establecido entre personas racionales. Ese acuerdo es hipotético pues supone un contexto premoral en el que se adoptan las reglas y prácticas morales. Pero las partes del contrato son reales, son individuos determinados que se distinguen por su capacidad, su situación y sus intereses. En la medida en que esos actores acepten limitar sus elecciones restringiendo la búsqueda de sus propios intereses, están reconociendo una distinción entre lo que les es lícito hacer y lo que no les es lícito hacer. Como personas racionales que comprenden la estructura de su interacción, las partes del acuerdo reconocen un lugar para la mutua restricción y también para la dimensión moral aplicada a sus asuntos.

Evidentemente hay que demostrar que hay una racionalidad contractual que puede aplicarse a la moral. Y ésa es la tarea de nuestra teoría. Aquí nuestro interés inmediato es relacionar la

idea de esa racionalidad con la presentación de las distinciones morales fundamentales. Ese no es un proceso mágico. La moral no puede emerger como un conejo de un sombrero vacío. Antes bien, surge de manera bastante sencilla de aplicar la concepción maximizadora de la racionalidad a ciertas estructuras de interacción. La restricción mutua acordada es la respuesta racional a esas estructuras. La razón anula el supuesto contra la moralidad.

En una teoría contractual, el elemento genuinamente problemático no es la introducción de la idea de la moralidad, sino el paso desde el acuerdo hipotético a la restricción moral real. Supongamos que cada persona se reconoce como una de las partes del acuerdo. Los principios que forman el objeto del convenio son aquellos que la persona ha aceptado *ex ante* al pactar con sus pares, mientras se encontraba entre ellos en un contexto inicialmente desprovisto de restricciones morales. ¿Por qué ese individuo tendría la necesidad de aceptar *ex post* en su actual situación, esos principios que limitan sus elecciones? Una teoría que defienda una moral basada en el acuerdo tiene que responder a esta pregunta.

Históricamente, la idea del contrato moral parece haberse originado entre los sofistas griegos. Glaucón perfila un enfoque basado en el acuerdo del origen de la justicia en *La República* de Platón, pero significativamente, se lo presenta a Sócrates para que éste lo refute en lugar de defenderlo.[15] Nuestra teoría de la moral cae en una tradición impopular, como lo confirma la identidad de su mayor defensor, Thomas Hobbes. Hobbes transforma las leyes de la naturaleza —que se basan en la médula del pensamiento moral estoico y cristiano medieval— en preceptos de la razón que exigen que cada persona, obrando en su propio interés, renuncie a una porción de la libertad, con lo cual busca su propia supervivencia y su propio bienestar, en tanto los otros actúen de la misma forma.[16] Pero ese acuerdo sólo permite que aparezcan las verdaderas restricciones a través de la eficacia del soberano político; desde el punto de vista de la teoría moral el paso crucial exige la intervención de un *deus ex machina*. Sin embargo en Hobbes encontramos el verdadero antecedente de la teoría de la moral que presentaremos aquí. Sólo recientemente esa posición comenzó a tener derivaciones

significativas. G.R. Grice ha desarrollado una explícita teoría del contrato y Kurt Baier ha reconocido las raíces hobbesianas que tiene su tesis central, según la cual "la verdadera *raison d'être* de una moralidad es producir razones que estén por encima de las razones del interés personal en aquellos casos en los que si cada uno busca su propio interés puede dañar a los demás". [17]

Al apuntalamiento conceptual que puede encontrarse en Hobbes, Grice y Baier, queremos añadir el rigor de la elección racional. Por supuesto, la teoría moral resultante no ha de ser necesariamente una teoría respaldada por esos autores. Pero el hecho de apelar a la elección racional nos permite establecer, con una nueva claridad y una nueva precisión, por qué las personas racionales se adherirían *ex ante* a principios constrictivos, qué características generales deben tener dichos principios como objetos de un acuerdo racional, y por qué las personas racionales estarían dispuestas a aceptar *ex post* las restricciones acordadas.

3.2 Una posición útil y ventajosa para apreciar la racionalidad de la restricción resulta de yuxtaponer dos ideas formuladas por John Rawls. Un defensor del contrato considera que la sociedad es "una empresa cooperativa que apunta a la ventaja mutua", emprendida por personas "concebidas como seres que no se preocupan por los intereses de los demás".[18] El defensor del contrato no pretende que todas las sociedades reales sean empresas cooperativas; no necesita que todas ellas ofrezcan la esperanza de una ventaja mutua. Antes bien, el defensor del contrato supone que, en general, una sociedad —considerada como un conjunto de instituciones, prácticas y relaciones— puede ofrecerle a cada persona el máximo beneficio que no podría esperarse en un "estado de naturaleza" no social y que sólo una sociedad semejante puede atraer la adhesión voluntaria de cada individuo racional. El partidario del contrato no necesita afirmar que las personas reales no se interesan por sus congéneres; en realidad suponemos que todos los seres humanos se caracterizan por cierto grado de sociabilidad. Pero el defensor del contrato considera que la sociabilidad es algo que enriquece la vida humana y para él, la sociabilidad se transforma en una fuente de explotación si induce a las personas a

aceptar instituciones y prácticas que, a no ser por los sentimientos de camaradería que esas personas experimentan, les resultarían costosas. Sin duda, el pensamiento feminista nos ha clarificado muy bien esta situación, quizá la forma más primaria de la explotación humana. Por lo tanto, el defensor del contrato insiste en afirmar que una sociedad no puede atraer la adhesión voluntaria de una persona racional si, sin apelar a sus sentimientos por los demás, no le ofrece una expectativa de beneficio neto.

Si las instituciones y prácticas sociales pueden beneficiar a todos, todos deberían aceptar que cierto conjunto de arreglos sociales constituyen una empresa cooperativa. La preocupación que cada persona tiene en satisfacer sus propios intereses, debería asegurar su voluntad de unirse a sus compañeros en esa empresa capaz de garantizarle una expectativa de mayor satisfacción. Por supuesto, esa persona puede rechazar cierta empresa propuesta por no considerarla suficientemente ventajosa, al evaluar tanto la distribución de los beneficios que ofrece como la disponibilidad de alternativas. Ofrecer un beneficio mutuo es una condición necesaria pero no suficiente para que se acepte un conjunto de arreglos sociales y se los considere una empresa cooperativa. Pero suponemos que un conjunto de arreglos que ofrezca el beneficio mutuo, también debe ser mutuamente aceptable; una teoría del contrato debe establecer las condiciones de suficiencia.

La racionalidad de un acuerdo sobre la sociedad concebida como una empresa cooperativa no parece constituir un problema. El paso que hay que dar para ir del acuerdo hipotético *ex ante* referente a un conjunto de arreglos sociales, a una adhesión *ex post* a tales arreglos parece completamente directo. Si una persona se ha unido voluntariamente a la empresa, ¿por qué querría no continuar con ella? ¿Por qué harían falta las coacciones?

Las prácticas e instituciones de una sociedad desempeñan un papel coordinador. Digamos, sin tratar de formular una definición precisa, que una práctica es coordinadora si cada persona prefiere adherirse a ella con la condición de que (la mayor parte de) los demás también lo hagan, pero no prefiere adherirse a ella salvo que (la mayor parte de) los demás no lo

hagan.[19] Y digamos que una práctica es beneficiosamente coordinadora si cada persona prefiere que las demás se adhieran a dicha práctica antes que adherirse a la no práctica y que no prefiere (fuertemente) que las demás se adhieran a alguna otra práctica alternativa. El ejemplo de Hume, de dos personas que reman en un bote que ninguna de las dos puede conducir sola, es un ejemplo muy sencillo de una práctica coordinadora beneficiosa.[20] Cada uno de los remeros prefiere remar si el otro también rema y prefiere no remar si el otro no lo hace. Y cada uno prefiere que el otro reme en lugar de que actúe de cualquier otra forma.

Vale la pena señalar que una práctica coordinadora no tiene por qué ser beneficiosa. Entre las personas pacíficas, que consideran que las armas sólo son un instrumento de defensa, cada uno puede preferir estar armado con la condición de que (la mayor parte de) los demás lo estén, y no estar armado sólo en el caso de que (la mayor parte de) los demás no lo esté. Estar armado es una práctica coordinadora, pero no es una práctica beneficiosa, pues cada uno prefiere que los demás no estén armados.

No hay que subestimar las ventajas coordinadoras de la sociedad. Pero no todas las prácticas sociales beneficiosas son coordinadoras. Digamos que una práctica es beneficiosa si cada persona prefiere que (casi) todos los demás se adhieran a dicha práctica antes que (la mayor parte de) los demás prefieran no realizar dicha práctica, y además no prefiere (fuertemente) que (casi) todos los demás se adhieran a alguna otra práctica. No obstante, puede darse el caso de que cada persona prefiera no adherirse a la práctica si (la mayor parte de) los demás lo hacen. En una comunidad en la cual los fondos recaudados por los impuestos se gastan de manera razonablemente sensata, cada persona puede preferir que casi todos paguen sus impuestos antes que preferir que no lo hagan y sin embargo, puede preferir no pagar sus propios impuestos mientras los demás lo hagan. Pues los pagos que cada persona hace contribuyen de manera insignificante a los beneficios que ella recibe. En una sociedad semejante, las personas pagarán voluntariamente sus impuestos sólo si cada una acepta cierta limitación en la búsqueda de sus intereses individuales; de lo contrario, cada una sólo pagará

sus impuestos si sufre la coerción ya sea de la opinión pública, ya sea de la autoridad pública.

La racionalidad de un acuerdo sobre la sociedad entendida como empresa cooperativa, no parece constituir todavía un problema. Pero el paso que hay que dar para ir del acuerdo hipotético *ex ante* sobre un conjunto de arreglos sociales, a la adhesión *ex post* ya no parece tan directo. Podemos entender por qué razón una persona que se une voluntariamente a la empresa no quiera continuar voluntariamente con ella. Cada uno se une a la empresa con la esperanza de beneficiarse con la adhesión de los demás, pero se abstiene con la esperanza de beneficiarse con su propia defección.

En los próximos dos capítulos ofreceremos un análisis del valor, la racionalidad y la interacción que nos brindará una formulación precisa de la cuestión que acabamos de identificar. Antes de internarnos en esa reflexión, deberíamos suponer que si cada persona elige su mejor curso de acción, el resultado debería ser el más ventajoso posible. Al llenar nuestros formularios de impuestos podríamos advertir, *inter alia*, que el beneficio individual y la ventaja mutua con frecuencia resultan intereses encontrados. Nuestra teoría desarrolla las implicaciones que puede tener el hecho de advertirlo y comienza por situar el conflicto entre el beneficio personal y la ventaja mutua dentro del marco de la elección racional.

3.3 Aunque una bien lograda teoría del contrato invalide el supuesto contra la moralidad que surge de su concepción de los individuos independientes y racionales, no puede dejar de considerar seriamente ese supuesto. Por consiguiente, la primera concepción central de nuestra teoría es que existe una zona moralmente libre, un contexto dentro del cual las restricciones de la moralidad no tendrían lugar. [21] La zona libre resulta ser ese terreno tan familiar a los economistas: el mercado perfectamente competitivo. Por supuesto, semejante mercado es una idealización; hasta qué punto puede llegar a ser real en una sociedad humana es una pregunta empírica que está más allá del objetivo de nuestra indagación. Nuestro argumento sostiene que en un mercado perfectamente competitivo, el beneficio mutuo queda asegurado por la actividad irrestricta de cada individuo que busca su propia satisfacción máxima, de modo tal

que, racionalmente, no hay lugar para la restricción. Además, puesto que en el mercado cada persona goza de la misma libertad en sus elecciones y acciones que tendría en un estado de aislamiento de sus semejantes, y puesto que el resultado del mercado refleja el ejercicio de la libertad de cada persona, no existe ninguna base que permita encontrar una parcialidad en las operaciones del mercado. De modo que, moralmente, tampoco hay lugar para la restricción. El mercado da un buen ejemplo de un ideal de interacción entre personas que, al no preocuparse por los intereses de los demás, sólo tienen que seguir los dictados de sus propios intereses individuales para participar efectivamente en una empresa destinada al beneficio mutuo. Y aquí no estamos hablando de una empresa *cooperativa* pues reservamos esa designación para las empresas que necesitan la natural armonía de cada uno con los demás, asegurada por la estructura de la interacción del mercado.

Por lo tanto, el mercado perfectamente competitivo es una lente a través de la cual la moralidad puede apreciarse más claramente. Si el mundo fuera un mercado semejante, la moral no sería necesaria. Pero denigrar el valor de la moralidad no permite lograr una armonía artificial allí donde no puede darse una armonía natural. El mercado y la moral comparten la conciliación no coercitiva del interés individual y el beneficio mutuo.

En el caso en que el beneficio mutuo exige la restricción individual, esa conciliación se logra mediante el acuerdo racional. Como ya hemos señalado, una condición necesaria para tal acuerdo es que su resultado sea mutuamente ventajoso. Nuestra tarea es pues establecer una condición suficiente. Este problema se remite a una parte de la teoría de los juegos, la teoría de la negociación racional, y se divide en dos cuestiones.[22] La primera es el problema de la negociación propiamente dicha que en su forma general está destinada a seleccionar un resultado específico —dada una cantidad de posibilidades mutuamente ventajosas— y una posición de negociación inicial. La segunda es pues determinar la posición de negociación inicial. El tratamiento de estas cuestiones aun tiene que alcanzar el consenso, de modo que desarrollaremos nuestra propia teoría de la negociación.

Al resolver el problema de la negociación se establece un principio que gobierna tanto el proceso como el contenido del acuerdo racional. Volveremos sobre esta cuestión en el capítulo 5 en el que presentaremos una medida de lo que cada persona pone en juego en una negociación: la diferencia entre lo mínimo que podría aceptar a cambio de no acordar y lo máximo que podría recibir en lugar de quedar excluido del acuerdo por voluntad de los demás. Y sostendremos que la idéntica racionalidad de los negociadores exige que la máxima concesión (medida como una proporción de lo que el que concede pone en juego) sea lo más pequeña posible. Y llamamos a esto el principio de la concesión relativa mínima. Que tiene su equivalente en la exigencia de que el menor beneficio relativo (medido nuevamente como una proporción de lo que cada uno pone en juego) sea lo mayor posible. De modo que formulamos un principio equivalente de beneficio relativo máximo que, según sostenemos, incluye las ideas de justicia e imparcialidad en una situación de negociación y por lo tanto sirve como la base de la justicia. Por consiguiente, la concesión relativa mínima o el beneficio relativo máximo constituye la segunda concepción central de nuestra teoría.

Si la sociedad debe ser una empresa cooperativa destinada a producir la ventaja mutua, sus instituciones y prácticas deben pues satisfacer o tratar de satisfacer ese principio. Porque si nuestra teoría de la negociación es correcta, la concesión relativa mínima gobierna pues el acuerdo *ex ante* que está en la base de una empresa cooperativa justa y racional. Pero en la medida en que los arreglos sociales restrinjan nuestras elecciones *ex post* reales, la cuestión del consentimiento exige particular atención. Supongamos que siempre es racional estar de acuerdo con las prácticas que aseguran el beneficio relativo máximo, ¿no es acaso igualmente racional desentenderse de esas prácticas si sirven al interés de uno? ¿Es racional internalizar los principios morales en las propias elecciones o sólo acatarlos en la medida en que las restricciones coercitivas externas refrenen nuestros propios intereses? La debilidad de la teoría tradicional del contrato ha sido su incapacidad de mostrar la racionalidad del acatamiento.

Aquí presentamos la tercera concepción central de nuestra

teoría: la maximización restringida. Distinguimos entre la persona que está dispuesta a maximizar directamente su satisfacción o a lograr su interés en las elecciones particulares que haga, y aquella otra persona que está dispuesta a acatar las restricciones morales mutuamente ventajosas, con la condición de que espere un acatamiento similar por parte de los demás. Esta última persona es lo que llamamos un maximizador restringido. Y los maximizadores restringidos, al mantener una interacción con los demás maximizadores restringidos, gozan de oportunidades de cooperación de las que carecen aquellos otros, identificados en primer término. Evidentemente, los maximizadores restringidos a veces, al estar dispuestos al acatamiento, pierden, pues pueden actuar de manera cooperativa en la errada esperanza de reciprocidad de los demás que, en cambio, se benefician a expensas de ellos. Sin embargo, mostraremos que, en condiciones plausibles, el beneficio neto que cosechan los maximizadores restringidos de la cooperación excede los beneficios de la explotación que los otros pueden esperar obtener. Partiendo de esta idea llegamos a la conclusión de que es racional disponerse a adoptar una conducta maximizadora restrictiva internalizando los principios morales para gobernar las propias elecciones. El defensor del contrato puede mostrar que es irracional admitir las apelaciones al interés contra el acatamiento de aquellos deberes basados en la ventaja mutua. [23]

Pero el consentimiento sólo tiene una base racional dentro del marco de una empresa plenamente cooperativa, en la cual cada participante establece una interacción voluntaria con sus pares. Y esto nos remite nuevamente a la segunda cuestión que mencionamos al hablar de la teoría de la negociación: la posición negociadora inicial. Si las personas están dispuestas a acatar el acuerdo que determina lo que toma cada cual de la mesa de negociación, deben pues considerar inicialmente aceptable lo que cada cual pone sobre la mesa de la negociación. Y si lo que alguno aporta incluye el fruto de una interacción previa forzada sobre los otros negociadores, ese aporte ya no resultará inicialmente aceptable. Si alguien se apodera del producto de mi trabajo y luego me dice: "Hagamos un trato", quizás yo esté obligado a aceptar, pero no lo haré voluntariamente.

Por consiguiente nos vemos obligados a restringir la posición negociadora inicial, mediante una estipulación que prohíba mejorar la posición de alguno de los negociadores mediante una interacción que empeore la posición de otro.[24] Ninguna persona debería estar en peores circunstancias en la posición negociadora inicial de lo que estaría en un contexto no social carente de interacción. Por consiguiente, la estipulación restringe la base sobre la cual se miden lo que cada persona pone en juego en el acuerdo y el beneficio y la concesión relativos de esa persona. Mostraremos luego que esa estipulación induce a una estructura de derechos personales y de propiedad personal que son básicos para establecer acuerdos sociales racional y moralmente aceptables.

La estipulación es la cuarta de las concepciones medulares de nuestra teoría. Aunque es parte de una moral basada en el acuerdo, no es el producto de un acuerdo racional. Antes bien es una condición que cada persona debe aceptar para hacer posible tal acuerdo. Entre seres que, aun siendo racionales, pueden no desear comprometerse recíprocamente en una empresa cooperativa destinada a lograr la ventaja mutua, la estipulación puede carecer de fuerza. Nuestra teoría le niega un lugar a la restricción racional y por lo tanto a la moralidad, fuera del contexto del beneficio mutuo. Una explicación contractual de la moral no da lugar a deberes cuyos efectos son estrictamente redistributivos, es decir que transfieren pero no aumentan los beneficios, ni a deberes que no impliquen reciprocidad por parte de los demás. Semejantes deberes ni tendrían base racional ni estarían apoyados por consideraciones de imparcialidad.

A las cuatro concepciones medulares, cuyo papel ya hemos esbozado, agregaremos una quinta: el punto arquimédico desde el cual un individuo puede modificar el mundo moral.[25] Para conferir semejante poder moral, el punto arquimédico tiene que asegurar una absoluta imparcialidad: la posición detrás del "velo de ignorancia" imaginada por John Rawls. En el capítulo 8 concluiremos la exposición de nuestra teoría moral relacionando la elección que hace una persona que ocupa el punto arquimédico con las demás ideas medulares. Mostraremos que la elección arquimédica está claramente concebida, no como un caso limitador de la decisión individual que se realiza en la

incertidumbre, sino antes bien como un caso limitador de la negociación. Y luego expondremos cómo cada una de nuestras ideas medulares (la estipulación contra el mejoramiento de uno mismo mediante el empeoramiento de los demás, la zona moralmente libre proporcionada por el mercado perfectamente competitivo, el principio de concesión relativa mínima y la disposición de la maximización restringida) puede estar relacionada directa o indirectamente, con la elección arquimédica. Al abrazar estas otras concepciones centrales de nuestra teoría, el punto arquimédico revela la coherencia de una moral basada en el acuerdo.

4 Una teoría contractual de la moral, desarrollada como parte de la teoría de la elección racional, tiene una fuerza evidente. Nos permite demostrar la racionalidad de las restricciones imparciales en la busca del interés individual aplicadas a las personas que no se interesan en los intereses de los demás. De modo que la moralidad puede apoyarse en una concepción débil y ampliamente aceptada de racionalidad práctica. Ninguna explicación alternativa de la moralidad logra ese cometido. Aquellos que declaran que los principios morales son objetos de una elección racional realizada en circunstancias especiales no pueden establecer la racionalidad del acatamiento real de esos principios. Y aquellos que declaran establecer la racionalidad de tal acatamiento apelan a una concepción fuerte y controvertida de la razón que parece incorporar supuestos morales previos. Ninguna explicación alternativa genera una moral que pueda considerarse una restricción racional a la elección y a la acción, partiendo de una base no moral o moralmente neutra.

Pero la teoría moral contractual parece tener además de esa fuerza una grave debilidad. Ya hemos observado que para un teórico contractual la moralidad exige un contexto de beneficio mutuo. John Locke sostenía que "un hobbesiano... no ha de admitir fácilmente muchos deberes claros de la moralidad".[26] Y esto parece poder aplicarse igualmente al más reciente sucesor de Hobbes. Nuestra teoría no implica ninguna preocupación fundamental por la imparcialidad, sino únicamente una preocupación derivada de los beneficios del acuerdo, y esos beneficios están determinados por los efectos que cada persona puede ejercer en los intereses de sus congéneres. Sólo los seres

cuyas aptitudes físicas y mentales son en general iguales o mutuamente complementarias pueden esperar que la cooperación resulte beneficiosa para todos. Los seres humanos se benefician de la interacción que mantienen con los caballos, pero no cooperan con los caballos y probablemente no los beneficien. Entre seres desiguales, una parte puede beneficiarse más ejerciendo coacción sobre la otra y, según nuestra teoría, no habría razón para que esa persona se refrenara. Podemos condenar todas las relaciones coercitivas, pero sólo dentro del contexto del beneficio mutuo, esa condenación puede apelar a una moral con bases racionales.

Las relaciones morales que existen entre los participantes de una empresa cooperativa destinada a obtener la ventaja mutua tienen una base firme en la racionalidad de los participantes. Ha sido plausible representar a la sociedad que surgió en la Europa occidental y en Norteamérica en los siglos recientes como una empresa semejante. Y ello se debe a que la sociedad occidental descubrió cómo poner coto a los esfuerzos de los individuos que trabajan para su propio bien adhiriéndolos a la causa del beneficio mutuo siempre creciente.[27] De ese descubrimiento resultaron no sólo una explosión de la cantidad de bienes materiales y del número de personas, sino también, y lo que es más importante, un aumento sin precedentes en el lapso promedio de vida y una ampliación, antes inimaginable, de la cantidad de ocupaciones y actividades efectivamente accesibles a la mayor parte de los individuos sobre la base de sus deseos y habilidades.[28] Con la ganancia personal vinculada con el progreso social, el individuo fue liberándose paulatinamente de los lazos coercitivos interpuestos por las costumbres y la educación, la ley y la religión, que caracterizaron a las sociedades anteriores. Pero al quitarle esos lazos al individuo, quizá se haya dado demasiado crédito a la eficacia de las instituciones que obran de modo semejante al mercado y se le haya prestado escasa atención a la necesidad de la interacción cooperativa que exige una restricción limitada pero real.[29] Una moral basada en el acuerdo expresa pues la real preocupación que cada uno de nosotros tiene por mantener las condiciones en las cuales la sociedad pueda ser verdaderamente una empresa cooperativa.

Pero, si bien las circunstancias que permitieron a las

personas considerarse recíprocamente como socios contribuyentes de una empresa conjunta pasaron por alto la crítica de Locke a la finalidad de la moralidad contractual, otras circunstancias pueden volver a ponerla en primer plano. Nuestra sociedad está pasando de una tecnología que permitió que una proporción siempre creciente de personas aumentara el nivel promedio de su bienestar a una tecnología —cuyo mejor ejemplo son los avances de la medicina— que permite una transferencia siempre creciente de los beneficios hacia las personas que disminuyen aquel promedio.[30] Tales personas no forman parte de las relaciones morales basadas en una teoría contractual.

Más allá de nuestra preocupación por el objetivo de las relaciones morales está la cuestión referente al lugar que ocupan dichas relaciones en una vida humana ideal. Glaucón le pidió a Sócrates que refutara una explicación contractual de la justicia, porque creía que semejante enfoque debía considerar la justicia como algo apreciable para las personas que son mutuamente dependientes, pero intrínsecamente inapreciable, de modo tal que "parece corresponder a la forma del trabajo penoso e inútil".[31] La cooperación es la segunda forma de interacción que exige concesiones y restricciones que cada persona puede preferir evitar. En realidad, cada uno tiene la secreta esperanza de poder ser injusto con éxito y cae fácilmente víctima de la más peligrosa de las vanidades, es decir, la que lo persuade de que en realidad él es superior a sus semejantes y por lo tanto puede ignorar impunemente los intereses de los demás en procura de los propios. Como dijo Glaucón aquel que "es verdaderamente un hombre" rechazará las restricciones morales.[32]

Una teoría contractual no contradice ese enfoque, puesto que deja completamente abierta la posibilidad de satisfacción de los deseos humanos, pero tampoco necesita de él. Antes bien, ¿no cabría suponer que la satisfacción de los seres humanos depende de una red de relaciones sociales cuya estructura misma los tienta constantemente a abusar de ella? Las restricciones de moralidad sirven pues para regular las relaciones sociales que se manifiestan incapaces de autorregularse. Y limitan los intereses que podamos tener por un ideal compartido de sociabilidad.

La cooperación puede entonces parecer una segunda forma

de interacción, no porque vaya contra nuestros deseos, sino porque cada persona podría preferir una armonía natural en la cual pudiera satisfacer sus deseos sin restricción. Pero una armonía natural sólo podría existir si nuestras preferencias y aptitudes se ajustaran unas a otras de modo tal que cada una impidiera el libre desarrollo de las demás. La armonía natural exigiría un nivel superior de artificio, la posibilidad de dar forma a nuestra naturaleza de modos que, por lo menos hasta que no se perfeccione la ingeniería genética, son imposibles; por lo demás, si fueran posibles, seguramente no serían deseables. Si la individualidad humana está destinada a florecer, debemos pues esperar cierto grado de conflicto entre los objetivos e intereses de las personas, antes que una armonía natural. El mercado y la moralidad suavizan ese conflicto al conciliar la individualidad con el beneficio mutuo.

En los últimos capítulos de nuestra investigación consideraremos qué puede decirse sobre esta interpretación del lugar que ocupan las relaciones morales en la vida humana. Para hacerlo tendremos que poner el acento en las cuestiones especulativas que subyacen más allá y por debajo de la teoría de la elección racional. Y podemos encontrarnos con una interpretación alternativa de lo que presentamos como una teoría de la moral.[33] Apuntamos a crear un vínculo entre la racionalidad de la maximización individual y la moralidad de la restricción imparcial. Supongamos que realmente encontramos ese vínculo. ¿Cómo debemos interpretar ese descubrimiento? ¿Son nuestras concepciones de racionalidad y moralidad —y por lo tanto también el vínculo contractual existente entre ellas—, como nosotros quisiéramos que sean, puntos fijos del desarrollo de un marco conceptual que nos permite formular verdades prácticas permanentes? ¿O sólo estamos contribuyendo a la historia de las ideas de una sociedad particular, en la cual peculiares circunstancias han alimentado una ideología de la individualidad y la interacción que resulta coherente con una moral basada en el acuerdo? ¿Estamos contando una historia sobre las ideas que podrá parecerles a nuestros descendientes tan extraño como la Forma del Bien o el Motor Inmóvil nos parecen a nosotros?

Notas

1 Véase David Hume, *An Enquiry concerning the Principles of Morals*, sec.ix, punto ii en L. A. Selby-Bigge (comp.), *Enquiries concerning Human Understanding and concerning the Principles of Morals*, 3ª. ed., Oxford, 1975, pág. 280.

2 Ogden Nash, "Kind of an Ode to Duty", *I Wouldn't Have Missed It: Selected Poems of Ogden Nash*, Boston, 1975, pág. 141.

3 Véase David Hume, *A Treatise of Human Nature*, Libro III Parte III, sec. III. L. A. Selby-Bigge (comp.), Oxford, 1888, págs. 413-418.

4 Por lo tanto uno puede proponer una teoría del error del lenguaje moral; sobre la idea de una teoría del error, véase J. L. Mackie, *Ethics: Inventing Right and Wrong*, Harmondsworth, Middx., 1977, capítulo 1, particularmente las páginas 35, 48-49.

5 La idea de que las apelaciones morales son persuasivas fue desarrollada por C. L. Stevenson; véase *Ethics and Language*, New Haven, 1944, especialmente los caps. 6 y 9.

6 Nuestro esbozo de la teoría racional le debe mucho a J.C. Harsanyi, véase "Advances in Understanding Rational Behavior", en *Essays on Ethics, Social Behavior, and Scientific Explanation*, Dordrecht, 1976, págs 89-98, y "Morality and the theory of rational behaviour", en A. Sen y B. Williams (comps.), *Utilitarianism and beyond*, Cambridge,1982, págs. 42-44.

7 J. Rawls, *A Theory of Justice*, Cambridge, Mass., 1971, pág.16; J. Harsanyi, "Morality and the theory of rational behaviour", pág. 42.

8 Véase *ibid*, pág. 12.

9 Véase *ibid*., pág. 11: "los principios...deben asignar derechos y deberes básicos y determinar la división de los beneficios sociales". Los principios destinados a los individuos se distinguen de los principios de justicia, véase pág. 108.

10 Véase la distinción que hace Rawls entre lo "razonable" y lo "racional" en "Kantian Constructivism in Moral Theory", *Journal of Philosophy* 77, 1980, pág. 528-530.

11 Véase Rawls, *A Theory of Justice*, págs. 14-15, y Harsanyi, "Morality and the theory of rational behaviour", págs. 44-46, 56-60.

12 Véase Harsanyi, "Morality and the theory of rational behaviour", pág. 62.

13 Esta concepción de la racionalidad práctica aparece con particular claridad en T. Nagel, *The Possibility of Altruism*, Oxford, 1970, especialmente, cap. 10. También puede encontrarse en la teoría moral de R. M. Hare; véase *Moral Thinking*, Oxford, 1981, espec. capítulos 5 y 6.

14 Véase Harsanyi, "Advances in Understanding Rational Behavior", pág. 89; también J. Elster, *Ulysses and the Sirens: Studies in rationality and irrationality*, Cambridge, 1979: "El enfoque de la conducta humana basado en la idea de la elección racional es, sin lugar a dudas, el modelo más aprovechable...", pág. 112.

15 Véase Platón, *República*, 358b-359b.

16 Thomas Hobbes, *Leviathan*, Londres, 1651, cap. 14, págs. 64-65.

17 Véase G .R. Grice, *The Grounds of Moral Judgement*, Cambridge, 1967, y K. Baier, *The Moral Point of View: A Rational Basis of Ethics*, Ithaca, Nueva York, 1958; las citas son de la pág. 309.

18 Rawls, *A Theory of Justice*, págs. 4 y 13.

19 Este análisis está relacionado con mi caracterización de una convención en "David Hume, Contractarian", *Philosophical Review* 88, 1979, págs. 5-8.

20 Véase Hume, *Treatise*, iii,ii,ii, pág.490.

21 Este es el tema del capítulo 4 presentado más adelante. Véase mi análisis anterior en "No Need for Morality: The case of the Competitive Market", *Philosophic Exchange* 3, n° 3, 1982, págs. 41-54.

22 Se encontrarán referencias a la bibliografía referente a la negociación racional, en las notas 12-14 del capítulo 5, *infra*.

23 Esta conclusión se apoya en una reinterpretación de la concepción maximizadora de la racionalidad que desarrollamos en el capítulo 6; véase sobre todo el párrafo inicial de la sec. 3.1.

24 Sobre la idea de la estipulación, véase nota 1 del capítulo 7.

25 Sobre la idea de un punto arquimédico, véase Rawls, *A Theory of Justice*, págs. 260-265.

26 Locke MS, citado en J. Dunn, *The Political Thought of John Locke*, Cambridge, 1969, págs. 218-219.

27 No ofrecemos una explicación para este descubrimiento. Parece que no hay ninguna razón para suponer que sea el resultado de una investigación deliberada.

28 Sobre el aumento del lapso promedio de vida, véase N. Eberstadt, "The Health Crisis in the USSR.", *New York Review of Books* 28, n° 2, 1981, pág. 23. En cuanto a la ampliación de la cantidad de ocupaciones accesibles obsérvese que " Todavía en 1815 tres cuartos de su población [de Europa] estaba empleada en el campo..." *The Times Concise Atlas of World History*, G. Barraclough (comp.), Londres, 1982, pág. 82.

29 De modo que la concepción de un hombre económico dedicado a apropiarse ilimitadamente de lo que está a su alcance llega a dominar el pensamiento social. Los efectos de esa concepción son uno de los temas de mi "The Social Contract as Ideology", *Philosophy and Public Affairs* 6, 1977 págs 130-164.

30 El problema aquí es no preocuparse por las personas de edad

que han pagado ya los beneficios que reciben con su actividad productiva anterior. Las terapias destinadas a prolongar la vida tienen sin embargo un potencial redistributivo ominoso. El problema primario es ocuparse de los incapacitados. Hablando eufemísticamente, se dice que hay que capacitarlos para llevar vidas productivas cuando los servicios exigidos superan cualquier producción posible, con lo cual se oculta una cuestión que, es comprensible, nadie quiere afrontar. Sin poner el acento en esas cuestiones, he tratado de comenzar a darle un enfoque contractual a ciertos aspectos de los sistemas de salud en "Unequal Need: A Problem of Equity in Access to Health Care", *Securing Access to Health Care: The Ethical Implications of Differences in the Availability of Health Services*, 3 vols., President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, Washington, 1983, vol. 2, págs 179-205.

31 *Platón, Republic*, 358a, trad. A. Bloom, Nueva York, 1968, pág. 36.

32 *Ibíd*, 359b, pág. 37.

33 Las ideas expresadas en este párrafo han recibido la influencia de R. Rorty, véase especialmente "Method and Morality", en Norma Haan, R. N. Bellah, P. Rabinow y W. M. Sullivan (comps.), *Social Science as Moral Inquiry*, Nueva York, 1983, págs. 155-176.

II

Elección: razón y valor

1. "Nada es bueno o malo, sólo el pensamiento le da ese carácter."[1] Pero si las cosas consideradas en sí mismas no son ni buenas ni malas, si realmente no existe un reino del valor independiente de los seres animados y de sus actividades, el pensamiento no es pues la actividad que da origen al valor. Hume nos recuerda que "la razón es y sólo debe ser la esclava de las pasiones", y si bien esa declaración de Hume ha sido ampliamente combatida, nosotros la defenderemos. [2] El deseo y no el pensamiento, la volición y no la cognición, son los impulsores del bien y del mal.

Nosotros podríamos querer pasar por alto las grandes cuestiones del valor analizadas por los filósofos occidentales desde la época de Sócrates y los sofistas. Pero las respuestas a las preguntas formuladas por ellos están implícitas en la teoría de la elección racional. Aunque presentar una completa teoría del valor es algo que está más allá del objetivo de la presente indagación, descubriremos sin embargo que, al discutir los conceptos básicos de la teoría de la elección —la preferencia, la utilidad, la maximización—, llegaremos no sólo a delinear un instrumento técnico sino también a reflexionar sobre la razón y el valor. En este capítulo analizaremos tanto el instrumento técnico como los supuestos necesarios para explicar la moralidad como algo que emerge del acuerdo racional.

Pero el acuerdo racional en sí mismo, y el papel que desempeña en la interacción, son cuestiones que analizaremos más adelante. Ahora pondremos el acento en la elección *paramétrica*, en la cual el actor considera que su conducta es la única variable de un ambiente fijo. En la elección paramétrica el actor se considera a sí mismo como el único centro de acción. En cambio, en la elección *estratégica* se incluye la interacción; en ella el actor considera que su conducta es sólo una variable entre muchas, de modo tal que su elección debe responder a las expectativas que él tenga sobre las elecciones que hagan los

demás que a su vez han de elegir respondiendo a sus propias expectativas.[3] En el caso de los ambientes paramétricos, la elección racional está claramente definida y nosotros sostendremos que la moralidad forma parte de la definición de elección racional en el caso de ciertos ambientes o contextos estratégicos. Pero ya nos estamos anticipando. ¿Cómo deberíamos entender la elección paramétrica?

En el capítulo I hablamos del interés o los intereses de cada persona y relacionamos la racionalidad con la satisfacción del interés individual y la moralidad con una restricción imparcial en la busqueda de ese interés. Pero el interés es un concepto que oscila incómodo entre la perspectiva del propio individuo y la del otro. La teoría de la elección racional considera primordial un concepto aun más claramente subjetivo y conductista que el interés, la relación de *preferencia individual.* La preferencia se relaciona con las situaciones; uno habla de preferir una manzana a una pera, pero si quisiéramos ser más estrictos tendríamos que hablar de que preferimos comer una manzana a comer una pera en cierto contexto o conjunto de contextos dados. Por supuesto la teoría de la elección racional se interesa primariamente por las preferencias entre situaciones, concebidas como posibilidades de acción alternativas y realizables. Por consiguiente estas situaciones no son objetos directos de elección, sino que más bien son resultados posibles de las acciones que uno puede elegir realizar. La teoría no analiza las relaciones particulares de la preferencia que se consideran datos últimos, sino que analiza conjuntos de esas relaciones, conjuntos que representan cada uno las preferencias de un individuo sobre los pares de resultados que pueden obtenerse en una situación de elección.

En su sentido más general, la racionalidad práctica se identifica con la maximización. Por lo tanto los problemas de la elección racional son de un tipo matemático bien conocido: uno apunta a maximizar cierta cantidad sujeta a cierta restricción. La cantidad que se busca maximizar debe estar vinculada con la preferencia; hemos hablado vagamente de la ventaja o el beneficio o la satisfacción, pero la teoría de la elección racional define una medida precisa de la preferencia, la *utilidad*, e identifica la racionalidad con la maximización de la utilidad. Por

lo tanto, se atribuye la utilidad a situaciones, consideradas éstas como objetos de relaciones de preferencia. La restricción bajo la cual ha de maximizarse la utilidad queda establecida por las posibilidades de acción. El actor racional maximiza su utilidad eligiendo entre un conjunto finito de acciones y considerando como resultados posibles cada una de las situaciones de un conjunto finito dado.

Alguien que quisiera objetar este enfoque podría estar de acuerdo con identificar la racionalidad práctica con la maximización, pero insistiría en que una medida de preferencia individual no representa una cantidad apropiada para maximizar. Es racional maximizar el *valor*; la teoría de la elección racional implícitamente identifica el valor con la utilidad, pero el impugnador se opone a esa identificación. Un opositor radical insistiría en afirmar que el valor no es una medida, sino una norma o una pauta de la preferencia. El valor no depende de la preferencia, sino que antes bien la preferencia racional depende del valor. Explicaremos nuestro rechazo a esta posición clarificando nuestra concepción del valor. Un opositor más moderado podría aceptar que el valor es una medida, pero una medida incapaz de medir las preferencias en bruto que podrían caracterizarse por la desinformación, la inexperiencia o la falta de consideración. Aceptaremos este punto de vista, en tanto éste se refiera al modo en que se sostienen las preferencias. Por lo tanto estableceremos un conjunto de condiciones necesarias para una *preferencia considerada* y que deben satisfacerse si ha de identificarse la utilidad —como una medida de preferencia— con el valor, y la maximización de la utilidad con la racionalidad.

También debemos establecer las condiciones formales necesarias para definir la utilidad. Hemos dicho ya que la utilidad es una medida de la preferencia; más precisamente es una medida de los resultados que representan las relaciones de preferencia. Si tenemos un conjunto dado de esas relaciones, que constituyen las preferencias de un individuo en una situación de elección, diremos que ha de atribuírsele una utilidad a cada resultado posible, de modo tal que uno pueda inferir de las utilidades la preferencia de la persona entre dos resultados cualesquiera. Entre dos resultados posibles cualesquiera habrá

de preferirse aquel que proporcione la mayor utilidad. Pero no todos los conjuntos de relaciones de preferencia permiten definir semejante medida. Supongamos, por ejemplo, que Bruce prefiere comer una manzana a comer una pera y comer una pera a comer un melocotón y también prefiere comer un melocotón a comer una manzana. En este caso la medida requerida de las preferencias de Bruce debe asignarle un valor mayor a comer una manzana que a comer una pera, a comer una pera que a comer un melocotón y a comer un melocotón antes que a comer una manzana. Pero esto es imposible. Las preferencias de Bruce violan las condiciones que debe cumplir un conjunto de relaciones de preferencia necesarias para definir una medida. Las preferencias que no pueden satisfacer esas condiciones son incoherentes; y nosotros debemos establecer las condiciones de la *preferencia coherente.*

Para evitar malentendidos, deberíamos señalar que si bien podemos inferir las preferencias de las utilidades, no podemos explicar las preferencias por las utilidades. La utilidad es simplemente una medida y su función es similar a la de la temperatura. Si la temperatura que se registra en Aix-en-Provence es superior a la que se registra en Toronto, podemos inferir que en Aix-en-Provence hace más calor que en Toronto. Pero no podemos explicar el mayor calor que hace en Aix atendiendo a su temperatura superior. Allí no hace más calor porque la temperatura registrada sea más alta, sino que más bien podemos decir que allí se registra una temperatura más alta, porque es un lugar donde hace más calor. Las medidas en sí mismas no cumplen un papel explicativo.

La utilidad se atribuye directamente a los posibles resultados, pero la elección se refiere a las acciones. Si puede conocerse con certeza el resultado de cada acción posible, la elección entre las acciones es pues equivalente a la elección entre sus resultados. El problema de maximización que debe resolverse al elegir racionalmente es entonces muy sencillo: uno selecciona aquella acción que le proporciona el resultado de mayor utilidad. (Si hay varias acciones que proporcionan idénticos resultados uno elige indiferentemente entre ellas). Basta con una medida *ordinal* de los resultados, es decir que la elección racional en situaciones de certeza sólo exige colocar los resultados en orden de jerarquía.

Y puesto que la mayor parte de los economistas clásicos analizan la elección en situaciones de certeza, la utilidad ordinal satisface las necesidades del economista.[4]

Pero en la mayor parte de las situaciones de la vida real, aun cuando el ambiente pueda considerarse fijo, sus características no pueden conocerse cabalmente. De modo que los resultados de ciertas elecciones son riesgosos o inciertos. (En la teoría de la elección racional el riesgo se asocia con la probabilidad objetiva y la incertidumbre con la subjetiva.)[5] Por lo tanto, la elección entre las acciones debe basarse no sólo en las utilidades de sus posibles resultados sino también en las probabilidades de esos resultados. El problema de la maximación es más complejo y en general sólo puede resolverse si se da una medida *en intervalos* de la preferencia. De las utilidades de cada uno de los resultados posibles de un conjunto dado debemos poder inferir no sólo el orden de preferencia sino también las fuerzas relativas de las preferencias. Si Bruce prefiere comer una manzana a comer una pera y comer una pera a comer un melocotón, de sus utilidades debemos pues poder inferir si su preferencia por comer una manzana en lugar de una pera es más fuerte que su preferencia por comer una pera en lugar de un melocotón.

Por consiguiente, las condiciones formales de un conjunto de preferencias necesarias para definir la utilidad se dividen en dos grupos: aquéllas suficientes para obtener una medida puramente ordinal, y aquéllas exigidas adicionalmente para obtener una medida de intervalos. Después de esbozar estos dos grupos de condiciones podremos extender nuestro análisis a solucionar el problema de la maximización desde el simple caso de la elección en una situación de certeza hasta una elección en situación de riesgo e incertidumbre. La elección estratégica que implica interacción plantea una serie de cuestiones adicionales a las que nos referiremos en el capítulo 3.

Si se cumplen las condiciones necesarias para las preferencias coherentes, podemos presentar una medida de los objetos de preferencia. Si también se cumplen las condiciones de la preferencia considerada, podemos identificar esa medida con el valor y su maximización con la racionalidad. En consecuencia, podemos hablar de las condiciones necesarias para una preferencia coherente y considerada como condiciones de una prefe-

rencia racional. Pero al hablar así no debemos dejarnos engañar, porque ninguna de las condiciones se refiere a la satisfacción de las preferencias particulares. El interés de la razón no son las preferencias particulares, sino el modo en que éstas se sostienen y sus interrelaciones. Una vez más estamos de acuerdo con Hume, en este caso cuando Hume dice que "no es contrario a la razón preferir que se destruya el mundo entero a un rasguño en mi dedo".[6] Lo que resultaría contrario a la razón sería sostener semejante preferencia de un modo desconsiderado o unirla a ciertas preferencias alternativas. Pero considerándola en sí misma, no podemos evaluar su racionalidad. Evidentemente, no deberíamos negar que semejante preferencia puede ser un síntoma de una perturbación emocional grave rayana en la insania. Pero ese carácter patológico debe establecerse considerando cómo se adquirió y cómo se mantuvo tal preferencia. De todos modos, sólo puede resultar confusión de la demasiado común suposición de que una perturbación de los sentimientos debe implicar una falla de la razón.

Las condiciones de la preferencia racional no exigen que las preferencias de las diferentes personas se adapten a un molde común. Las preferencias de cada persona determinan los valores de esa persona de manera completamente independiente de los valores de los demás. De modo tal que una determinada situación se caracteriza no por un valor único sino por un conjunto de valores, uno correspondiente a cada persona, en los cuales se incluyen o pueden incluirse las preferencias. El valor proporciona no una medida uniforme y única de la preferencia sino una medida relativa para cada evaluador. Y aunque los valores se atribuyan a las situaciones, esa atribución tiene que ver con actitudes y no con observaciones, es subjetiva y no objetiva. Como medida de la preferencia, el valor depende y debe depender de las preferencias para tener una existencia real. Por lo tanto, la teoría de la elección racional considera el valor como una medida subjetiva y relativa y no como una norma objetiva y absoluta.[7] Estos contrastes son controvertidos y exigen un comentario adicional.

La teoría de la elección racional también considera la razón práctica como algo estrictamente instrumental. Esto no queda completamente implícito al identificar la racionalidad con la

maximización, porque, como hemos observado, uno podría suponer que la cantidad que ha de maximizarse no es una medida sino una norma de la preferencia, una norma objetivamente inherente a las situaciones cuya comprensión implica un ejercicio de la razón. En una perspectiva semejante, la razón podría no ser meramente instrumental sino que también podría referirse a los fines de la acción. Pero, al identificar la racionalidad con la maximización de una medida de preferencia, la teoría de la elección racional repudia todo interés por los fines de la acción. Los fines pueden inferirse de las preferencias individuales; si las relaciones entre esas preferencias y el modo en que se las sostiene cumplen con las condiciones de la elección racional, la teoría acepta pues cualquier fin que ellos impliquen.

Por consiguiente, una concepción instrumental de la racionalidad se vincula con la idea de identificar el valor con la utilidad. Si el valor fuera una norma, la razón podría pues desempeñar un papel en la determinación de esa norma que podría ir más allá de la mera maximización. Y si el valor no fuera ni una norma ni una medida, la razón no desempeñaría ningún papel práctico, no habría nada que maximizar. La teoría de la elección racional se desenvuelve entre el dogmatismo que significa asumir una norma para la preferencia y el escepticismo que implica negar una medida de preferencia. Puesto que nuestro objetivo es lograr una elección racional que se haga cargo de "remolcar" la teoría moral, deberíamos asegurarnos que este análisis impida el naufragio.

Hemos planteado tres temas que han de ocupar el resto de este capítulo. En la próxima sección analizaremos la preferencia considerada. Luego discutiremos la preferencia coherente y en la última sección defenderemos la concepción que considera el valor como una medida subjetiva y relativa.

2.1 Hasta ahora hemos hablado de si es posible señalar una diferencia entre la preferencia y la elección, no sólo desde el punto de vista conceptual sino también desde el punto de vista operativo. Hemos supuesto que las preferencias que tiene una persona sobre todos los pares de resultados posibles pueden considerarse datos básicos, que es posible definir una medida de esos resultados y que de esa medida pueden inferirse las preferencias de dicha persona, y que, como base para la elección,

puede presentarse una regla que maximice esa medida. Puesto que hemos considerado que la preferencia y la elección son conceptos diferentes, la racionalidad de las elecciones puede estimarse determinando si esas elecciones maximizan la satisfacción de la preferencia. Pero, como bien sabe el lector familiarizado con la economía, éste no es el modo que utiliza el economista para relacionar la preferencia y la elección; y consecuentemente, no es el modo en que entiende la maximización y la utilidad.[8] Para el economista, la elección sólo es accesible a la observación. La racionalidad es asumida dondequiera que sea posible, de modo tal que se considera que una persona se comporta racionalmente si puede darle a sus elecciones una interpretación racional. El valor se identifica, pues, con lo que sea que esa persona maximice. Las preferencias no se dan como datos, sino que se infieren; la preferencias de una persona pueden identificarse con cada una de las relaciones de un conjunto dado entre los resultados posibles tales que su medida quede maximizada por las elecciones de dicha persona. En la perspectiva del economista, la preferencia se manifiesta en la elección y no tiene una significación operativa independiente.

Sea o no adecuado para la economía, este enfoque no basta para elaborar una teoría moral. El economista supone que una persona se comporta de manera irracional o no logra maximizar el valor si y sólo si no se le puede dar una interpretación maximizadora a su conducta. Aunque la razón y el valor desempeñen un limitado papel normativo en cuanto a prescribir y criticar una elección, ese papel se expresa en el único mandado: "¡Maximiza!" Decir que uno debe maximizar la utilidad como una medida de la preferencia, no agrega nada, puesto que la utilidad se identifica sencillamente con aquello que permite que se interprete como maximizadora la conducta de una determinada persona.[9]

Si la razón y el valor están destinados a desempeñar papeles normativos en el esquema necesario para comprender la acción humana suministrado por la teoría de la elección racional, debe pues ser posible dar la orden de maximizar satisfacciones adicionales. Si una persona puede comportarse de un modo maximizador y sin embargo no totalmente racional, podríamos decir que la maximización, aunque necesaria, no

puede ser suficiente para la racionalidad. Pero la preferencia manifestada es una concepción muy pobre para poder satisfacer esa satisfacción adicional. Si la única manera que tenemos de conocer las preferencias de una persona es a través de sus elecciones, cualquier cosa que esa persona maximice debe ser la medida de aquellas preferencias. Para pasar de la explicación economicista a una perspectiva de elección racional que permita comprender la conducta racional, debemos comenzar por reconsiderar el concepto de preferencia.

Sin embargo, no debemos suponer que el acento que ponen los economistas en la preferencia manifestada es simplemente un error. Ciertamente, las elecciones ponen de manifiesto las preferencias; una concepción de la preferencia que considerara el vínculo existente entre la preferencia y la elección como algo meramente accidental, sería aun más evidentemente inadecuada que una que sostenga que la preferencia parasita en la elección. Lo que debemos rechazar del enfoque economicista es la insistencia en afirmar que la elección nos brinda el único acceso a la preferencia. Debemos hacer una distinción entre una dimensión de conducta de la preferencia puesta de manifiesto en la elección y una dimensión de actitud, expresada en palabras. Al elegir una manzana de una frutera que contiene manzanas y peras una persona pone de manifiesto una preferencia: comer una manzana en lugar de una pera. Al decir que a uno le gustan más las manzanas que las peras uno expresa una preferencia por comer una manzana en lugar de una pera. Por supuesto, en algunas situaciones la elección es verbal, uno pone de manifiesto sus preferencias expresándolas. Pero en otros contextos, las dos dimensiones de la preferencias son completamente distintas y pueden estar en conflicto. Karen expresa una preferencia por leer obras filosóficas en lugar de ver un programa cómico por la televisión, pero pasa su tiempo libre sentada ante el televisor mientras su biblioteca de filosofía acumula polvo.

Ahora estamos en posición de agregar algún contenido, aunque todavía no sea en absoluto suficiente, al mandato de maximizar. La utilidad es una medida de la preferencia, pero no de la preferencia puesta de manifiesto o de actitud, tomada aisladamente de la otra. Es una medida de ambas dimensiones en tanto ambas concuerden. Si las dos dimensiones están en

conflicto, aun cuando sea posible definir una medida para cada una, ninguna puede identificarse con la utilidad. Si la preferencia puesta de manifiesto y la preferencia expresada por una persona divergen, esa persona tiene pues los valores confundidos y carece de una base adecuada para hacer una elección racional. Una condición necesaria para suponer que una persona maximiza el valor y por lo tanto actúa racionalmente, es que puedan interpretarse sus elecciones como actos que maximizan la medida de sus preferencias expresadas.

Carecer de una medida común a la conducta y a las actitudes es un signo de irracionalidad. Pero no podemos atribuir sin más rodeos esa irracionalidad a las elecciones de una persona porque no coincidan con sus actitudes, ni tampoco a sus actitudes porque éstas no coincidan con sus elecciones. En circunstancias específicas, este tipo de atribución puede estar claramente justificada. Pero en general sólo podemos llegar a la conclusión de que la divergencia entre la elección y la actitud indica irracionalidad, pero no podemos situar esa irracionalidad de manera más precisa. Porque tal irracionalidad podría quedar superada tanto si las elecciones de la persona llegaran a coincidir con lo que expresa, como si sus actitudes llegaran a coincidir con sus elecciones.

En muchas situaciones una persona elige sin expresar verbalmente una preferencia. En otras situaciones una persona expresa una preferencia sin estar en posición de elegir. Evidentemente, no suponemos que en esas situaciones la persona sea irracional sólo porque no se puedan medir las dos dimensiones de su preferencia. Antes bien suponemos que en tanto le sea posible un individuo ha de confirmar verbalmente su elección o ha de confirmar con su conducta su actitud. Suponemos que existe una coincidencia entre la preferencia puesta de manifiesto y la preferencia expresada, salvo que tengamos una prueba clara de divergencia. Suponemos que hay una única medida de preferencia, salvo que nos encontremos ante medidas opuestas. Aceptamos el supuesto del economista que sostiene que debe atribuirse racionalidad a la conducta de una persona en ausencia de una prueba que indique lo contrario, pero rechazamos las limitaciones extremas sobre la prueba que el economista permite.

2.2 La elección puede estar errada de muchos modos. La

reina Gertrude elige la copa envenenada; una joven dice "sí" (ante una propuesta o una proposición) y luego, aunque sólo inmediatamente después, se da cuenta de que ha cometido un error. Un turista recién llegado a Francia pide una botella Nuits-Saint-Georges para acompañar su *sole meunière*. Podemos decir que la elección de Gertrude careció de información, la de la joven de consideración y la del turista de experiencia.

La elección de Gertrude, aunque fatalmente errada, no fue irracional. La reina no tenía razón para creer que la copa estaba envenenada; quería beber y bebió lo que creyó que podía beber. Si hubiese tenido una razón para sospechar de la copa y puesto que su objetivo no era el suicidio, su elección habría sido irracional. Pero no necesariamente debemos atribuir esa irracionalidad a sus preferencias; sus creencias podrían haber estado equivocadas. Su elección podría haber reflejado su incapacidad para evaluar correctamente la prueba referente a sus circunstancias.

Podría suponerse que los valores de una persona deben identificarse con la medida de sus preferencias informadas. Pero nosotros rechazamos esta idea. En cambio, relacionamos las preferencias de una persona con su situación de creencia o información. Gertrude elige el veneno pero no cree que lo que eligió sea veneno y su elección no revela una preferencia por el veneno. Al considerar que sus preferencias están en relación con sus creencias, no tenemos razón para considerarlas erradas o para negarnos a identificar su medida con sus valores.

En algunos casos, podemos inferir lo que habría preferido una persona si hubiese estado correcta o completamente informada sobre sus circunstancias. Gertrude habría preferido una copa de vino. Pero en otros casos, nos es completamente imposible determinar qué habrían preferido las personas puesto que éstas no tienen preferencias formadas relacionadas con los hechos que les tocan vivir en determinado momento. Podemos relacionar la preferencia que tenían los aztecas por los sacrificios humanos con su creencia de que los dioses estaban sedientos de sangre humana, pero, ¿cómo podrían haberse alterado esas preferencias si los aztecas no hubiesen tenido esas creencias? No lo sabemos ni tenemos ningún modo de saberlo. Las preferencias de los aztecas se formaron en el contexto de sus creencias;

más allá de ese contexto, los aztecas no tenían, o no les hacía falta tener, preferencias. Suponer que cada persona debe tener un conjunto de preferencias "ciertas" o "verdaderas" relacionadas con sus circunstancias reales, aunque con frecuencia encubiertas por creencias falsas o incompletas, no parece psicológicamente plausible.

Un análisis completo del papel que le cabe a la creencia en relación con la elección y la preferencia nos alejaría demasiado del presente objetivo. Aceptamos el esquema explicativo general: existiendo una creencia, la elección maximiza la satisfacción de la preferencia. Pero rechazamos la trivialización que resulta de negar el acceso independiente mediante una prueba a cada uno de sus términos: el acceso a la elección, a la preferencia y a la creencia. La elección puede extraviarse a causa de una creencia errada: la elección puede ser irracional porque la persona se formó una creencia irracional. Pero esos errores de la creencia no afectan la identificación del valor con la medida de la preferencia ni la racionalidad de la elección basada en la preferencia.

Pasemos de la reina Gertrude al turista inexperto. Esperemos que éste antes de abandonar Francia haya descubierto que no prefiere acompañar el pescado con vino tinto. Su elección inicial reflejó que el hombre no estaba familiarizado con la cocina francesa; no tenía una preferencia firme o clara. Quizá su elección fue irracional; hubiese sido mejor que consultara al camarero o que estudiara minuciosamente la sección dedicada a las comidas y las bebidas de su guía, pero los camareros intimidan y las guías abruman. Podría ser que su elección reflejara sus preferencias y sus creencias lo mejor que él podía determinarlas en sus circunstancias. Pero no deberíamos considerar que su preferencia puesta de manifiesto inicialmente por el vino tinto con pescado sea una base para determinar sus valores, puesto que evidentemente el hombre carece de la experiencia necesaria para formarse una preferencia considerada.

El hecho de que la joven acepte una propuesta (o una proposición) tampoco nos da una base para determinar sus valores. Quizás ella no carezca de la experiencia necesaria para formarse una preferencia considerada, pero no reflexionó lo

suficiente antes de aceptar. Cuando reflexiona se da cuenta de que su preferencia firme era decir:"No".¿Cambió la joven su preferencia inicial o la corrigió? Ninguna de las dos alternativas es aceptable; la primera implicaría que sus preferencias sucesivas se basan en fundamentos idénticos; la segunda implicaría que inicialmente la joven falseó su preferencia. Una explicación mejor sería que, sin la debida consideración, la joven se forma una preferencia tentativa que su aceptación pone de manifiesto, y que luego la revisa. Cualquier elección pone de manifiesto una preferencia (de conducta), pero no todas las preferencias tienen la misma categoría, de modo tal que el rechazo de una preferencia tentativa no puede considerarse lo mismo que la alteración de una preferencia firme o considerada. Por supuesto, no todo rechazo de una preferencia tentativa implica necesariamente la formación de una preferencia firme. Al "no" de la muchacha podría haber seguido otro "sí"; pero podemos suponer que en este caso, una mayor reflexión sólo confirmaría la segunda preferencia. Y así consideramos que estos datos nos suministran la base para determinar los valores de la joven.

Cuando no existe una cabal reflexión sobre las preferencias, la elección tiende a extraviarse. Pero, insistimos, no por ello se puede afirmar que la elección sea irracional. Posiblemente las circunstancias no den lugar a una adecuada reflexión; el asaltante que dice: "La bolsa o la vida" probablemente no pueda cumplir mi exigencia de tomarse tiempo para reflexionar. Y a veces, aunque el tiempo lo permita, la inversión de esfuerzo que implica la reflexión puede ser improductiva. Por ejemplo, si en los estantes del supermercado me encuentro ante catorce marcas diferentes de jabón para lavar la ropa, simplemente tomo la caja más cercana. Pero si estoy decidiendo dónde pasar mis vacaciones estaría obrando irracionalmente si no me convenciera de que una reflexión adicional podría alterar mi preferencia.

Si uno elige sin tener una experiencia adecuada de los resultados posibles, más tarde podría desear echarse atrás en dicha elección. Si uno elige sin reflexionar lo suficiente sobre los resultados, más tarde puede desear alterar su elección, en caso de que eso sea posible. En ninguno de los dos casos, uno supondría que la elección original fue positiva o que maximizó la medida de las propias preferencias consideradas. Si a pesar

de todo la persona considera que eligió racionalmente, debería considerar que su elección se basó en la mejor estimación que pudo hacer de sus preferencias consideradas en la situación en que se encontraba, es decir, las preferencias que podrían pasar la prueba de la reflexión y la de la experiencia.

Pero las preferencias consideradas, aunque estables, no tiene por qué ser inmutables. Mark, de quince años, prefiere escuchar música disco a escuchar a Bach; veinte años después Mark prefiere escuchar a Bach. Podemos juzgar que su preferencia por la música disco indica una falta de sensibilidad y discernimiento musicales. Pero a medida que va adquiriendo sensibilidad, Mark cambia su preferencia y llega a considerar superior a la anterior su preferencia actual. Llega a pensar que es mejor preferir escuchar música de Bach. Pero, ¿esto no implica acaso que cuando originalmente prefería escuchar música disco su preferencia era inconsiderada?

La falta de sensibilidad musical del Mark de quince años podría invalidar cualquier juicio musical que hiciera, pero no sus preferencias musicales. El hecho de que haya llegado a preferir escuchar a Bach no le brinda una real preferencia por Bach que esté en conflicto con una preferencia meramente aparente por la música disco. Su preferencia por la música disco no es ni irreflexiva ni inexperta; escuchar música disco no le hace desear echarse atrás en su elección y si se viera obligado a escuchar a Bach, eso no le haría desear haberlo hecho voluntariamente. Más adelante, pasados varios años, Mark desea haber escuchado más a Bach en su juventud. Pero nada de sus preferencias primeras podría haber hecho de ésa una elección racional. Aunque una persona pueda lamentar sus elecciones pasadas, porque éstas no lograron maximizar la satisfacción de sus preferencias consideradas, esa persona también puede llegar a lamentarlas porque las preferencias cabalmente consideradas cuya satisfacción maximizaron esas elecciones no se corresponden con la preferencias actuales.

Un detractor podría señalar que nuestra argumentación ignora el juicio de Mark según el cual es mejor preferir escuchar a Bach. Mark lamenta sus elecciones pasadas, no solamente porque sus preferencias hayan cambiado, sino también porque ahora considera inferiores sus preferencias pasadas. Y las

elecciones basadas en preferencias inferiores están erradas. Pero, si bien en cierto sentido esta declaración es cierta, no es aplicable a las elecciones pasadas de Mark. Suponer que una preferencia es superior a otra es sencillamente preferir una a otra. Si una persona se encuentra en posición de elegir entre sus preferencias, racionalmente elegirá aquellas que considere superiores. Elegir las preferencias inferiores sería contradecir las preferencias expresadas mediante las preferencias puestas de manifiesto. Pero Mark no hace eso. A los quince años no consideraba inferior su preferencia por escuchar música disco y, de cualquier modo, no le importaba elegir entre las posibles preferencias musicales.

Una persona elige racionalmente al esforzarse por maximizar la satisfacción de aquellas preferencias que sostiene de un modo considerado en la situación de elección. Que esto pueda incluir preferencias que posteriormente uno puede desear no haber tenido, es un rasgo perturbador de la vida humana, pero no de una elección irracional. Igualmente perturbador para la vida es el hecho de que uno pueda tener preferencias que al mismo tiempo desearía no tener. Algunas personas prefieren fumar (y esa preferencia puede estar cabalmente considerada y hasta excesivamente experimentada) y sin embargo preferirían no preferir fumar. Puede ocurrir que fumar, aunque evidentemente satisface la preferencia por fumar, disminuya la posibilidad de que esa persona algún día llegue a preferir no fumar. Esa persona afronta una difícil elección, pero no es una elección que le plantee un problema a nuestra explicación de la elección racional. La fuerza de sus preferencias y el efecto que pueda ejercer esa elección en la satisfacción de tales preferencias determina qué considera esa persona racional hacer. Pero seguramente al lector han de ocurrírsele nuevas objeciones. ¿No estamos acaso pasando por alto cuestiones tales como los intereses y la prudencia? Por lo tanto volveremos sobre esas cuestiones y otras que se vinculan con ellas.

2.3 Las preferencias son consideradas si y sólo si no hay un conflicto entre su dimensión de conducta y su dimensión de actitud y si permanecen estables aun cuando se las somete a la experiencia y la reflexión. El valor es la medida de la preferencia considerada y la elección racional implica el empeño por

maximizar el valor. Cualquiera de entre varias posiciones similares puede oponerse a este enfoque. Una persona puede preferir fumar otro cigarrillo a prescindir de él, pero puede ser irracional elegir fumar porque fumar daña la salud y esa persona tiene *interés* en ser saludable, más allá de lo que prefiera. Otra persona puede preferir perseguir a su amor perdido, en lugar de buscar consuelo en otra parte, pero estaría sosteniendo de manera irracional esa preferencia si continuara una búsqueda desesperanzada que sólo le agregaría aflicción en lugar de permitirle alcanzar una mayor *felicidad* o una mayor *satisfacción* que si podría lograr dirigiendo sus esfuerzos hacia otro objeto amoroso. Otra persona puede preferir gastar todos sus ahorros en satisfacer sus deseos actuales en lugar de guardar algo de dinero para la vejez, pero sería una forma irracional de sostener esa preferencia maximizar la medida de las preferencias actuales en lugar de preocuparse *prudentemente* por maximizar la medida de sus preferencias consideradas en el curso de su vida futura.[10]

No podemos analizar exhaustivamente las objeciones que podrían presentársele a nuestra posición; sencillamente esbozaremos las razones principales que tenemos para rechazar los tres puntos de vista mencionados. En cada caso, la aparente credibilidad de la alternativa depende de cotejarla con el enfoque económico que sostiene que una elección es racional cuando maximiza la preferencia puesta de manifiesto. Una vez que uno presenta la preferencia de actitud o hasta la preferencia reflexiva como uno de los fundamentos de la elección racional, la credibilidad de la alternativa se desvanece.

Una persona que reconoce tener interés en no fumar muestra irracionalidad si sencillamente continúa fumando, sin hacer un esfuerzo por detenerse ni expresar una preferencia considerada más fuerte por continuar haciéndolo. Pero no tenemos necesidad de explicar esa irracionalidad cotejando el interés con la preferencia como una base para realizar la elección. Reconocer el interés por no fumar es expresar una preferencia que vale indistintamente en aquellas situaciones que fomentan el interés como en aquellas que no lo fomentan. Si la conducta electiva ignora ese reconocimiento, evidentemente hay una divergencia entre la preferencia expresada y la preferencia puesta de manifiesto.

Pero, ¿qué ocurre con la persona que no reconoce un interés por no fumar? Los defensores de la idea de que la elección racional se basa en el interés afirmarán que con todo esa persona tiene un interés en no fumar y elige irracionalmente ignorando ese interés. Pero aquí no hay un conflicto abierto entre las dimensiones de conducta y de actitud de la preferencia. Entonces, ¿no nos comprometemos con la idea de que la conducta del fumador es racional?

No necesariamente. La preferencia por fumar que revela esa persona no puede resultar contradecida por un interés reconocido por no fumar, porque la persona no reflexionó lo suficiente. Esa falta de reflexión puede ser inadvertida o deliberada puesto que la persona puede saber que si reflexiona tiene que reconocer un interés por no fumar. Si podemos suponer que la preferencia revelada o puesta de manifiesto por la persona no soportaría una reflexión que la contraríe, podríamos convenir pues en que la elección de ese fumador no es plenamente racional.

Pero ahora deberíamos preguntarnos qué ocurre con la persona que teniendo una amplia experiencia y reflexionando acabadamente no reconoce un interés por no fumar. ¿Qué ocurre con la persona que rechaza firmemente cualquier preocupación práctica por su salud, tanto en las palabras como en los hechos? Un interés reconocido es una preferencia expresada, pero desde nuestro punto de vista, un interés rechazado no puede tomarse como base para determinar ni los valores de una persona ni la racionalidad de sus elecciones.

Por lo tanto, aquí estamos en desacuerdo con el defensor de la racionalidad de la elección basada en el interés. Creer que la base de una elección es un interés no reconocido y en realidad rechazado en lugar de una preferencia considerada es en efecto considerar que la satisfacción de la preferencia está sujeta a la evaluación racional. Y esto exige una concepción objetiva del valor. Si consideramos la propuesta de basar la elección en el interés como una alternativa subjetivista de nuestro punto de vista, veremos que es una alternativa errada. Porque si se acepta que los valores son subjetivos, no hay pues ninguna base para la apelación más allá de lo que una persona reconoce, dando por sentado que esa persona reflexionó lo suficiente y

tiene una basta experiencia. Por supuesto que esa persona puede sostener sus creencias de manera irracional, pero eso no impugna sus valores, aunque sí puede demostrar que son inaplicables o irrelevantes en sus reales circunstancias.

Hay pocas personas que no reconocen que la salud sea uno de sus intereses y aun más pocas que no lo reconocerían después de contar con una adecuada experiencia y de dedicarle cierta reflexión. Para todos, salvo para esos pocos, sería en realidad irracional comportarse como si la salud no fuera algo que hay que tener en cuenta. Aquellos que creen que fumar es dañino para la salud, si son racionales, deben elegir no fumar, a menos que alguna preferencia más fuerte e igualmente considerada sobrepase el interés que tienen por su salud. Para llegar a esta conclusión no necesitamos considerar que el interés es la base de la elección racional y no nos parece que se pueda llegar a ninguna conclusión más profunda.

Por supuesto puede ocurrir que una persona llegue a considerar de interés algo que antes rechazó y puede entonces lamentar haberlo ignorado antes. Este reconocimiento puede reflejar un cambio de la creencia o de la preferencia o puede ser el resultado de la experiencia vivida por esa persona al ignorar ese interés. En este último caso, las elecciones pasadas de la persona no dieron un buen resultado. A causa de su inexperiencia, la persona fue incapaz de determinar correctamente sus preferencias consideradas. Si esa incapacidad indica o no irracionalidad debe depender, por supuesto, de las circunstancias particulares.

La segunda posición que debemos considerar brevemente es que la elección racional se basa en la satisfacción o la felicidad, una medida de la calidad de la experiencia, o el goce o el carácter hedonista. No existe un término particular que baste para expresar ese concepto que algunos proponen como la base de la elección, quizá porque precisamente no está muy claro qué es lo que ha de medirse. Pero lo esencial es que haya una medida, distinta de la utilidad y por lo tanto independiente de la preferencia, aplicable a las situaciones en virtud de la forma en que se las experimenta. El defensor de la felicidad supone que la elección racional tiende a maximizar esa medida.

Una dificultad inmediata que plantea esta opinión es que en realidad no existe una medida de nuestra experiencia, cuyos

títulos podamos evaluar. Pero supongamos que exista tal medida y que lo que se mida sea el *goce*, palabra que en este análisis consideraremos un término semitécnico. De modo que la satisfacción o la felicidad quedarían definidas como la medida del goce, del mismo modo en que la utilidad es la medida de la preferencia o la temperatura es la medida del calor.

Suponemos que la gente llega a conocer el goce y su medida, de modo tal que al formarse sus preferencias tiene en cuenta la satisfacción que espera le proporcionen los resultados posibles de sus elecciones. Ahora bien, hay dos posibilidades. Una es que las preferencias plenamente consideradas de la persona no se centren exclusivamente en el goce, de modo tal que la utilidad y la satisfacción no siempre se correspondan. Este constituiría seguramente el mejor ejemplo contra la idea de igualar la racionalidad con la maximización de la satisfacción. Si después de una plena reflexión la persona no reconoce el goce como el objeto apropiado de preferencia, ningún argumento puede pues mostrar que la maximización de su medida sea la exigencia primaria de la razón práctica.

Por otra parte, puede ocurrir que las preferencias plenamente consideradas se refieran exclusivamente al goce, de modo tal que la medida del goce esperado siempre se corresponda con la medida de la preferencia. Si fuera así, tendríamos que haber hecho un profundo descubrimiento psicológico sobre la preferencia. Tendríamos que haber descubierto que la preferencia tiene un único objeto: el goce. Y tendríamos que llegar a la conclusión de que es posible descubrir la preferencia real que tiene una persona entre dos situaciones cualesquiera sin apelar ni a sus elecciones ni a sus expresiones de preferencia, sin siquiera necesitar que la persona conozca ni tenga conciencia de las situaciones, con tal que podamos determinar cuánto goce obtiene en cada una de ellas. Y hasta podríamos explicar las preferencias de una persona atendiendo al goce y a su medida, de modo que podríamos decir que prefirió un resultado a otro, porque el primero le proporcionaba mayor satisfacción.

Si todo esto fuera así, tendríamos que aceptar que la satisfacción proporciona la base para realizar una elección racional. Pero esto podría no oponerse a nuestra posición actual ya que la satisfacción suministraría la base para hacer una

elección racional, como la medida del objeto de preferencia, el goce. Pero, por supuesto, no tenemos razón para creer que la preferencia tenga un único objeto mensurable, tal como habíamos supuesto que era el goce. Parece claro que no pueden correlacionarse las preferencias con cualquier dimensión o característica única de las situaciones a las que se refieren tales preferencias. En realidad, parece claro que las preferencias no dependen únicamente de las calidades de experiencia, como sostiene el defensor de la satisfacción o la felicidad.[11] Creemos que estas alternativa no merece un análisis adicional.

La tercera posición que se opone a nuestro enfoque de la elección racional basada en la preferencia identifica la racionalidad práctica con la prudencia. El defensor de la prudencia afirma que no es racional maximizar simplemente la satisfacción de las preferencias presentes de una persona —por más consideradas que ellas puedan estar— si esa persona no tiene en cuenta las preferencias que puede llegar a tener en el futuro. Evidentemente, uno no puede tener pleno conocimiento de su futuro, pero puede formarse expectativas razonables que permiten diferentes alternativas dentro de las variadas probabilidades. La persona prudente piensa en el mañana y así, afirma nuestro detractor, elige racionalmente.

Un ejemplo ha de ilustrar convenientemente en qué difieren las opiniones que basan la elección racional en la preferencia y en la prudencia. El invierno llegará y cuando llegue Susan preferirá que su casa cuente con un buen sistema de aislación. Para Susan, ¿no es racional encargar ahora los trabajos, cuando los precios son más bajos y se dispone fácilmente de los materiales y la mano de obra? Ese sería el curso de acción prudente, pero como aún es verano y Susan se halla tomando baños de sol junto a la piscina, ninguna preocupación referente al invierno la perturba. Susan sabe que más adelante se va a preocupar, pero por ahora no lo hace. El defensor de la racionalidad de la prudencia insiste en afirmar que la presente despreocupación de Susan es irracional. Pero, desde nuestro punto de vista, sólo una consideración que forme parte de las preferencias presentes de una persona puede suministrarle una base racional para hacer una elección. Si la despreocupación de Susan ha sido plenamente considerada, ella no tiene pues ninguna razón para encargar los trabajos de aislación ahora.

Aquí nos encontramos con el proverbial contraste de la cigarra y la hormiga: ¿no estamos acaso defendiendo la racionalidad de la negligente cigarra? En realidad sí lo estamos haciendo, en la medida en que la cigarra es reflexivamente negligente, pero no por ello estamos impugnando la racionalidad de la prudente hormiga. Maximizar sobre la base de las preferencias actuales no significa necesariamente ignorar las preferencias futuras. Uno puede interesarse ahora en el bienestar futuro y preferir una vida satisfactoria a una gratificación más inmediata. Pero también puede preferir lo contrario. Desde nuestro punto de vista, la prudencia es racional para aquellos que tienen una preferencia considerada por ser prudentes, pero no para aquellos que después de reflexionar plenamente no la tienen.

Nuestro desacuerdo con el defensor de la prudencia no estriba en si deben tenerse o no en cuenta las preferencias futuras, sino en el modo en que éstas se consideren. Ambos coincidimos en sostener que la persona irreflexivamente imprudente, que no piensa en el mañana, elige irracionalmente. Las preferencias consideradas de una persona por los posibles resultados de sus elecciones deben tener en cuenta los efectos que han de ejercer esos resultados en la persona. Pero una persona puede preferir ignorar esos efectos al obrar; una persona puede elegir no pensar en el mañana. Y en nuestra opinión esa imprudencia reflexiva no es irracional. El defensor de la prudencia afirma que la elección racional debe apuntar a la satisfacción máxima de todas nuestras preferencias, presentes y futuras, en la medida en que seamos capaces de determinar cuáles son. En nuestra perspectiva, la elección racional debe enderezarse a lograr la máxima satisfacción de nuestras preferencias consideradas presentes y nos referimos a una consideración que se extiende a todos los efectos futuros que tengamos la posibilidad de prever.

Existe un estrecho vínculo entre el defensor de la elección basada en el interés y el defensor de la prudencia. Ambos consideran el futuro de un modo temporariamente neutral y niegan que en la deliberación racional sobre lo que conviene hacer haya una prioridad por el momento presente. Nuestra posición rechaza esa neutralidad temporal, salvo en el caso en

que forme parte de la posición presente de un determinado individuo. En otras palabras, suponemos que la medida en que un individuo se identifica prácticamente con sus futuros sí mismos está determinada por sus preferencias presentes y que esas preferencias, como todas las demás, no están sujetas a una evaluación racional que atienda a la satisfacción de esas preferencias. Sin duda, la persona marcadamente imprudente, entregada a la gratificación instantánea, puede descubrir al probar el amargo sabor de sus elecciones, que lamenta su pasado insensato. Pero su irracionalidad es una cuestión que depende de sus preferencias consideradas y no una cuestión determinada conceptualmente por la pretendida atemporalidad de la razón. La razón práctica se coloca en el punto de vista del presente.[12]

Los defensores del interés, del goce y de la prudencia puede oponer objeciones plausibles contra la opinión de que la mera preferencia de conducta constituye una base suficiente para realizar una elección racional y para determinar el valor. Pero la opinión de que la elección racional se basa en la preferencia considerada —que no muestra divergencia entre la conducta y la actitud y manifiesta adecuada experiencia y reflexión— soporta las críticas de esas posiciones alternativas. Sin embargo, la preferencia considerada es sólo una parte de nuestra historia. La maximización en sí misma le impone condiciones a la preferencia. Pero esas condiciones, necesarias para definir una medida de intervalos, son completamente independientes de la forma que se sostienen las preferencias. Las preferencias inconsideradas pueden cumplir —y las consideradas quizá no— las condiciones necesarias para la maximización. Para simplificar nuestro análisis, de ahora en adelante hablaremos únicamente de preferencia y daremos por sentado que a menos que tengamos que establecer específicamente una diferencia, nos estaremos refiriendo a la preferencia considerada, y al hablar de la maximización de la utilidad nos estaremos refiriendo a la medida de la preferencia considerada.

3.1 Las condiciones formales necesarias para definir una medida de intervalos de la preferencia han sido extensamente analizadas en la bibliografía dedicada a la elección racional.[13] Nada podemos agregar al tratamiento matemático. En esta sección nuestro aspiración es meramente ofrecer una explicación no técnica de la preferencia coherente. Como ya observamos anteriormente nuestro análisis se divide en dos partes: primero bosquejamos las condiciones suficientes que debe cumplir una medida meramente ordinal de la preferencia; luego presentamos las condiciones adicionales necesarias para determinar una medida de intervalos.

Entre los miembros de un conjunto dado es posible definir una medida ordinal basándose en cualquier relación que ordene de algún modo a esos miembros. La relación coloca a los miembros en un orden: del primero al último, del mejor al peor, del máximo al mínimo y pueden darse los empates. La altura ordena de algún modo a los seres humanos del más alto al más bajo. La liga de béisbol ordena a los jugadores por el mejor promedio de bateo vitalicio desde el más efectivo (Ty Cobb) al peor (?). Puede definirse una medida ordinal de la preferencia entre los posibles resultados de cualquier situación de elección, si y sólo si la preferencia ordena de algún modo esos resultados, desde el más preferido al menos preferido. Y también aquí pueden darse los empates (preferencias iguales o indiferencia).

La primera condición necesaria para que una relación ordene de algún modo a los miembros de un conjunto es que dos miembros cualesquiera puedan ser comparados sobre la base de esa relación. Por lo tanto, en una situación de elección, dados dos resultados posibles cualesquiera, el elector debe o bien preferir uno al otro, o bien elegir de manera indiferente entre ellos. Esta es la condición de *completitud* que excluye los resultados no comparables desde el punto de vista de la preferencia.

Como condición de la preferencia de conducta, la completitud es trivial. Si una persona debe elegir entre dos resultados posibles, su elección pondrá pues de manifiesto una preferencia entre ellos. Pero, por supuesto, esa persona puede no desear o no poder expresar o reconocer una preferencia. En cambio, como condición de una preferencia considerada, la completitud no es trivial. El hecho de que no se cumpla no impide la elección, sólo impide la elección racional. Si una persona considera que dos resultados posibles no son comparables desde el punto de vista de la preferencia, es incapaz de elegir racionalmente entre ellos. Esta es una consecuencia directa de la opinión que sostiene que la elección racional se basa en la preferencia.

Pero debe advertirse que la necesidad de una condición de

completitud aparece porque la elección racional se basa en una relación existente entre los objetos de elección. Sea cual fuere esa relación, debe satisfacer tal condición. Si sustituyéramos la preferencia por el interés o el goce como bases de una elección racional, deberíamos poder comparar dos posibles resultados cualesquiera atendiendo a la satisfacción del interés o atendiendo al goce, a fin de elegir racionalmente entre ellos.

Gran parte del análisis que desarrollamos en esta sección se refiere, no tanto a la preferencia en sí misma, como a las condiciones que debe cumplir cualquier enfoque que identifique la racionalidad con la maximización de una medida de la relación que exista entre los posibles resultados de las acciones. Pero, por supuesto, debemos preguntarnos si es plausible imponer esas condiciones a la preferencia; si no podemos hacerlo, debemos buscar alguna otra relación sobre la base de la elección racional. De lo contrario, debemos desistir de la idea de identificar la racionalidad con la maximización. Suponemos que es plausible exigir que la preferencia satisfaga la completitud, que es plausible afirmar que una persona incapaz de comparar dos situaciones desde el punto de vista de la preferencia es, en realidad, incapaz de elegir racionalmente entre ellas.

La segunda condición que debe cumplir una relación para ordenar de algún modo a los miembros de un conjunto dado es que sea transitiva. Ilustraremos la condición *transitiva* aplicada a la preferencia con un sencillo ejemplo. Supongamos que entre los posibles resultados de la elección de Bruce estén comer una manzana, comer una pera y comer un melocotón. Si Bruce prefiere comer una manzana a comer una pera y comer una pera a comer un melocotón, la relación tiene un carácter transitivo si Bruce prefiere comer una manzana a comer un melocotón. Si Bruce come indiferentemente una manzana o una pera y come indiferentemente una pera o un melocotón, la relación tiene un carácter transitivo si Bruce come indiferentemente una manzana o un melocotón. Además, si Bruce prefiere comer una manzana a comer una pera y come indiferentemente una pera o un melocotón, o come indiferentemente una manzana o una pera y prefiere comer una pera a comer un melocotón, la relación es transitiva si Bruce prefiere comer una manzana a comer un melocotón.

Aun más que la completitud, el carácter transitivo representa una norma de la preferencia antes que una caracterización de la conducta y las actitudes reales. Sin dudas, en algunos contextos la norma se cumple. Si los posibles resultados son sumas de dinero podemos pues esperar un carácter perfectamente transitivo de la preferencia: siempre se preferirá una suma mayor a una menor y dos sumas iguales serán resultados indiferentes. Pero cuando no hay una dimensión cuantiva según la cual puedan ordenarse los resultados y cuando esa dimensión es el único objeto de interés de la elección entre los posibles resultados, es dable esperar que el carácter tránsitivo no se cumpla en el mundo real. Esto resulta aun más evidente en el caso de la indiferencia. Una persona es indiferente entre una primera manzana y una segunda, entre una segunda y una tercera...entre una decimonovena y una vigésima, sin embargo no tiene duda de que prefiere la primera a la vigésima. Cada pequeña diferencia entre las sucesivas manzanas desaparece detrás del umbral de discriminación de esa persona, pero si se suman todas las diferencias el resultado ha de ser claramente advertible. Los objetos o las situaciones suficientemente similares se consideran iguales, pero la suficiente similaridad no constituye una relación transitiva.

La imposibilidad de satisfacer una relación transitiva conduce a un círculo de la preferencia que podemos ejemplificar con la referencia a Bruce que hicimos en la primera sección de este capítulo. Bruce prefiere comer una manzana a comer una pera, comer una pera a comer un melocotón y comer un melocotón a comer una manzana. Si se manifiestan en la conducta, tales ciclos podrían despojar a la elección de toda racionalidad. Supongamos que Bruce tiene una manzana; antes de que la coma le ofrecemos un melocotón. Bruce prefiere el melocotón y acepta el cambio. Antes de que coma el melocotón le ofrecemos una pera; como Bruce prefiere la pera acepta el cambio. Pero antes de que Bruce coma su pera le ofrecemos cambiársela por su manzana original; como Bruce prefiere la manzana acepta el trato. Después de tres transacciones, en cada una de las cuales mejoró su situación desde el punto de vista de sus preferencias, Bruce se encuentra en el punto de partida.

En realidad, la situación de Bruce es aún peor de lo que

sugiere este ejemplo. Si Bruce prefiere un melocotón a una manzana, presumiblemente está dispuesto a ofrecer más de una manzana a cambio de un melocotón. Supongamos que está dispuesto a ofrecer 1 centavo. Y supongamos que está dispuesto a ofrecer el melocotón y 1 centavo por una pera y la pera y 1 centavo por la manzana. En ese caso, después de tres transacciones, en cada una de las cuales mejoró su situación desde el punto de vista de sus preferencias, Bruce se encuentra con su manzana original pero tres centavos más pobre. Evidentemente un comerciante astuto podría obtener ventaja de las preferencias intransitivas de Bruce y arruinarlo mediante una interminable rueda de transacciones semejantes. Y queda claro que Bruce (que como resultado de esas transacciones está cada vez en peor posición a pesar de que cada una de ellas mejoró su posición desde el punto de vista de sus preferencias) queda envuelto en una secuencia irracional de elecciones.

Cualquier relación para poder constituir un fundamento de la elección racional debe carecer de carácter cíclico o debe evitar los ciclos completos de la preferencia. El carácter plenamente transitivo es una exigencia más fuerte, necesaria para que pueda identificarse la racionalidad práctica con la maximización. Pero el hecho de que sea o no plausible no estriba únicamente en esa identificación. Si la única exigencia fuera el carácter acíclico, Bruce podría mostrarse indiferente entre comer una manzana y comer un melocotón, aun cuando prefiera comer una manzana a comer una pera y comer una pera a comer un melocotón. Al excluir tales combinaciones de preferencia, el carácter transitivo seguramente expresa nuestro ideal preteórico de la racionalidad aplicada a la elección.

La completitud y el carácter transitivo juntos son suficientes para asegurar que la relación de preferencia ordena de algún modo cualquier conjunto de resultados posibles. Cada uno de esos conjunto incluye pues un resultado preferido a todos los demás, o varios resultados mutuamente indiferentes pero cada uno preferido a todos los demás. Esos son los resultados más preferidos; la utilidad, la medida de la preferencia, queda maximizada mediante la selección del resultado más preferido.

Pero, como ya hemos observado antes, los objetos directos de la elección son las acciones y no los resultados. En situaciones

de certeza, en las cuales a cada acción corresponde un resultado definido, esta cuestión no plantea ningún problema desde el punto de vista de la maximización. Pero en condiciones de riesgo o de incertidumbre, en las cuales a cada acción corresponde un conjunto de posibles resultados, cada uno con su probabilidad (objetiva o subjetiva), puede resultar imposible elegir directamente un resultado más preferido. Supongamos que Bruce prefiere comer una manzana a comer una pera y comer una pera a comer un melocotón y que sus preferencias son transitivas. Hay dos fruteras, una con dos manzanas, cinco peras y un melocotón y la otra con cuatro manzanas, una pera y tres melocotones. Bruce tiene autorización para elegir una frutera de la cual recibirá luego una fruta seleccionada al azar. ¿Qué frutera debería elegir? Su orden de preferencias no le da una respuesta a esa pregunta puesto que no revela qué fuerza tiene su preferencia por las manzanas en lugar de las peras en comparación con su preferencia por las peras en lugar de los melocotones. Si Bruce prefiere intensamente las peras a los melocotones y sólo se inclina levemente por las manzanas en lugar de las peras, al elegir la primera frutera hará la mejor elección. Pero si prefiere intensamente las manzanas a las peras y sólo se inclina levemente por las peras en lugar de los melocotones, hará la mejor elección optando por la segunda frutera. Para determinar las fuerzas respectivas de las preferencias de Bruce es necesaria la medida de intervalos.

3.2 En situaciones de riesgo o incertidumbre, cada acción sometida a la elección de una persona se asocia a un conjunto de resultados. A cada resultado del conjunto se le asigna una probabilidad; en situaciones de riesgo, como ya observamos antes, esas probabilidades tienen una base objetiva, mientras que en las situaciones de incertidumbre, las probabilidades quedan determinadas por las estimaciones subjetivas del individuo. A los fines del presente análisis, puede ignorarse la distinción entre la probabilidad objetiva y la probabilidad subjetiva siempre que establezcamos que las probabilidades subjetivas deben asignarse de modo tal que se sumen a la unidad. Consideramos cierto que si se elige una acción ha de obtenerse uno y sólo uno de los resultados del conjunto asociado a ella.

Podemos imaginar que cada acción es como una lotería en la que los resultados asociados son los premios. Pero aquí tenemos que hacer una advertencia sobre esa terminología que es la utilizada habitualmente en la bibliografía referente a la elección racional. Las preferencias que cada actor tiene por cualquier elección considerada como una lotería deben depender únicamente de sus preferencias por los resultados o premios y de sus probabilidades. La lotería en sí misma es transparente a la preferencia; cada persona es indiferente a la "jugada" que implica la lotería. En las loterías de la vida real los participantes se interesan no sólo por los premios y sus probabilidades, sino también por el placer (o el desagrado) que conlleva hacer una apuesta. Pero en las loterías asociadas al riesgo y la incertidumbre, hacer una apuesta no es uno de los factores que afectan la preferencia.

La utilidad, la medida de la preferencia, queda determinada por una función de valor real, es decir, una función que les asigna números reales a los resultados de modo tal que, si a un resultado X se le asigna un número mayor que a un resultado Y, el resultado X ha de ser preferido al resultado Y, y si se les ha asignado el mismo número a los resultados X e Y, éstos son indiferentes. En las situaciones de certeza sólo las preferencias por los resultados participan de las elecciones de una persona, puesto que cada acción está asociada a un único resultado. Pero en las situaciones de incertidumbre y riesgo, las preferencias por las loterías de las que dependen los resultados también deben participar de la elección. Para que exista una medida de esas preferencias, tal que nos permita interpretar que las elecciones de la persona la maximizan, es necesario que además de la condición de completitud y el carácter transitivo se cumplan dos condiciones adicionales más sobre la preferencia. Si se cumplen esas condiciones, es posible definir una medida de intervalos sobre los resultados y las loterías que pueden darse en una particular situación de elección, tal que la utilidad asignada a cualquier lotería es su *utilidad esperada*. Esto es, la suma de los productos obtenidos al multiplicar la utilidad de cada uno de los premios (o resultados) de la lotería por su probabilidad.

El siguiente ejemplo puede ser de utilidad. Supongamos

que Bruce vaya a recibir una fruta tomada al azar de la primera de las fruteras mencionadas al final de la última subsección, es decir, la frutera que contiene dos manzanas, cinco peras y un melocotón. La probabilidad de que Bruce reciba una manzana es de 0,25, que reciba una pera es de 0,625 y que reciba un melocotón de 0,125. Si para Bruce la utilidad de una manzana es 1, la de una pera es 0,5 y la de un melocotón 0, la utilidad esperada de recibir un fruta tomada al azar de la frutera es [(1 x 0,25) + (0,5 x 0,625) + (0 x 0,125)], o 0,5625.

Decir que la utilidad de una lotería es su utilidad esperada equivale a decir que las preferencias sobre las loterías pueden inferirse de las utilidades esperadas de las loterías. Si la lotería X tiene una utilidad esperada mayor que la lotería Y, la lotería X ha de ser pues preferida a la lotería Y. Si consideramos la utilidad como una medida de intervalos de las preferencias que tiene Bruce sobre los resultados posibles, podemos inferir pues sus preferencias sobre las loterías que tienen como premios esos resultados, calculando sus utilidades esperadas.

La teoría de la elección racional identifica la racionalidad práctica con la maximización de la utilidad. Si la utilidad de una lotería es su utilidad esperada, en condiciones de riesgo o incertidumbre la racionalidad práctica se identifica con la *maximización de la utilidad esperada*. Esta es la tesis central de la teoría de la decisión bayesiana.[14] Y aunque esa teoría no atrae el beneplácito universal, debo afirmar (dogmáticamente y sin ponerme a analizar sus méritos y sus defectos) que no conocemos ninguna alternativa satisfactoria.

¿Cuáles son esas condiciones adicionales que deben exigírsele a la preferencia para obtener una medida de intervalos? La primera es la *monotonicidad*. Consideremos dos loterías, X e Y, que difieren entre sí en un solo aspecto: uno de los premios (o resultados) de X, digamos P, en Y está reemplazado por otro premio Q (y ambos tienen, por supuesto, la misma probabilidad). La condición de monotonicidad exige que si Q es preferido a P, Y debe ser preferido a X.

Veamos un ejemplo sencillo. Diana ha de recibir una fruta tomada al azar de una frutera. Inicialmente la frutera contiene tres manzanas, tres peras y cuatro melocotones. Pero antes de que Diana reciba su fruta, Mark toma uno de los melocotones y

lo reemplaza por una manzana. Si Diana prefiere recibir una manzana a recibir un melocotón, la monotonicidad exige pues que Diana prefiera la última lotería a la primera.

La segunda condición es la *continuidad*. Consideremos tres resultados posibles X, Y y Z, tales que X es preferido a Y e Y es preferido a Z. La condición de continuidad establece que debe haber una y sólo una lotería, cuyos premios sean X y Z, en la que Y sea indiferente. En otras palabras, que haya una única probabilidad p tal que la lotería $[pX + (1\text{-}p)Z]$ sea indiferente a Y.

De modo que si las preferencias de un individuo satisfacen las condiciones de completitud, de transitividad, de monotonicidad y de continuidad, hay una forma fácil (en principio) de asignar las utilidades que miden sus preferencias sobre los miembros de cualquier conjunto de resultados posibles, tales que la utilidad de cualquier lotería que tenga como premios los miembros de ese conjunto, sea su utilidad esperada. Supongamos que se le asigne una utilidad de 1 a los resultados preferidos y que X sea uno de esos resultados. Supongamos que se le asigne una utilidad 0 a los resultados menos preferidos y que Z sea uno de esos resultados. En ese caso, la utilidad de cualquier otro resultado Y es p, en la cual Y es indiferente a la lotería $[pX + (1\text{-}p)Z]$. Quizás el lector haya observado ya que esta asignación no es la única; si U es una función de utilidad que proporciona una medida de intervalos de las preferencias que un individuo tiene entre los miembros de un conjunto de resultados posibles, cualquier transformación lineal positiva de U es por ello también una función semejante. Pero aquí no necesitamos internarnos en estas complicaciones.

Aún no hemos comentado la plausibilidad de la monotonicidad ni de la continuidad. En efecto, la monotonicidad descarta las preferencias referentes a las loterías en sí mismas —las preferencias por los riesgos o contra los riesgos— consideradas como opuestas a las preferencias por los premios (o resultados posibles). Y al comenzar esta subsección dijimos que esto era algo que debía cumplirse; sin embargo, el lector puede objetar que en muchas situaciones esa exigencia es irrazonable. Si yo prefiero que mañana haya sol a que llueva, para que se cumpla la monotonicidad debo pues preferir una remota probabilidad de que haya sol (contra grandes probabilidades de que llueva) a saber con seguridad que va a llover. Pero seguramente —y es completamente razonable— yo puedo no desear eso, pues una remota probabilidad de que sea un día soleado sólo complicaría los planes que puedo hacer para mañana, en comparación con la certeza de que va a llover. Mis preferencias pueden ser perfectamente coherentes aun cuando yo prefiera la seguridad de que haya sol a la seguridad de que llueva, y la seguridad de que llueva a un 10 % de probabilidades de que salga el sol.

Del mismo modo puede sospecharse de la condición de continuidad. Tomaré prestado un ejemplo de una de las fuentes más difundidas (Luce y Raiffa): prefiero 1 dólar a 1 centavo y 1 centavo a la muerte. Si se cumpliera la condición de continuidad debería pues haber una lotería cuyos premios fueran 1 dólar y la muerte y en la cual 1 centavo resultara indiferente. Pero, ¿correría yo un riesgo suplementario de muerte a cambio de 1 dólar si no estuviese obligado a hacerlo? Si la respuesta es no, yo debería preferir 1 centavo a cualquier lotería cuyos premios fueran 1 dólar y la muerte.[15]

Hay respuestas aparentemente sencillas a las objeciones presentadas en estos ejemplos. La primera, como habrá podido observarse, refleja una mala comprensión de la condición de monotonicidad y de las características especiales que tienen las loterías a las cuales se aplica. Supongamos que soy indiferente entre pasar las vacaciones del año próximo en Grecia o en México. Si soy una persona racional, ¿debo ser igualmente indiferente entre la seguridad de cualquiera de los dos destinos y una lotería que tenga como premios las vacaciones en Grecia y las vacaciones en México? Si la lotería fuera a decidirse inmediatamente, es plausible sostener que la respuesta sea afirmativa. Pero si la lotería ha de decidirse un día antes de que comiencen las vacaciones, la respuesta bien puede ser no; no me importa mucho si he de pasar mis vacaciones en Grecia o en México, pero quiero saber ahora adónde iré, a fin de poder ajustar mis planes. En el segundo caso, la lotería incluye un período de riesgo epistémico que en el primero está ausente y que no necesariamente me resulta indiferente.

La monotonicidad impone la indiferencia ante el riesgo o la incertidumbre como tales. Pero no exige que haya indiferencia

entre la presencia y la ausencia de un período temporal de incertidumbre. La monotonicidad es una exigencia plausible para que haya una preferencia coherente sobre las loterías en las que los premios reales o los resultados se determinan inmediatamente, de modo tal que el riesgo asociado a la lotería no afecte el estado epistémico de los participantes. Pero las loterías de la vida real no necesariamente se deciden en el instante. Aunque yo prefiera que mañana salga el sol y no que llueva, puedo preferir la certeza de la lluvia a una probabilidad remota de que salga el sol, porque la lotería del tiempo no ha de decidirse hasta mañana con lo cual me deja en la subjetivamente indeseable posición de no saber qué esperar. Tales loterías de la vida real no excluyen la monotonicidad como una condición que debe cumplir la preferencia coherente pero revelan que hay ciertas limitaciones en la aplicación de esa condición.

El segundo ejemplo, que indica la falta de cumplimiento de la condición de continuidad, quizá se comprenda mejor si se limitan los contextos en los cuales se espera que se cumpla esa condición, es decir, exceptuando los casos extremos. Pero, ¿existe algún modo preciso de distinguir los extremos? ¿Estamos apelando a algo que no sea la respuesta visceral, pues generalmente la condición de continuidad se cumple pero a veces nuestras preferencias se ordenan de un modo jerárquico o lexicográfico? Sin responder a estas preguntas y sin plantear otros problemas relacionados con la monotonicidad, ampliamente representados en la bibliografía referente a la teoría de la utilidad, nos apartaremos de nuestro esbozo de la preferencia coherente (conscientes de que la teoría de la elección racional aún continúa siendo un territorio subdesarrollado) hacia otras preocupaciones igualmente complejas pero menos técnicas.

4.1 Presentar la utilidad como una medida de preferencia es en sí mismo un ejercicio intelectual que no determina el papel que debe cumplir la utilidad. Hemos afirmado que al relacionar la utilidad con la preferencia considerada, podemos tratarla como la norma de la elección e identificar la elección racional con la maximización de la utilidad. Esto equivale a igualar la utilidad con el valor. Los economistas han observado con agrado esto y han continuado su estudio sobre la toma de decisiones. Los filósofos, en cambio, deben hacer una pausa a fin de

examinar la concepción del valor que implica igualarlo con la utilidad.

Según esta concepción, el valor es una medida de la preferencia. Para poder existir una medida depende de lo que mida —si no hay preferencia, no hay valor—. Por supuesto, puede atribuirse el valor a un objeto —a una manzana, por ejemplo— o más propiamente, a una situación que incluya un objeto —comer una manzana—, sin suponer por ello que el objeto o la situación relacionada con él forme parte de las preferencias de cualquier persona. La manzana puede ser suficientemente similar a otras que estén incluidas entre las preferencias de la persona; lo cual nos permite inferir las preferencias hipotéticas de alguna persona referentes a las situaciones en las cuales el objeto puede estar incluido. Y de este modo puede asignarse el valor a esas situaciones. Pero sólo puede atribuírsele un valor a los objetos o situaciones en la medida en que, directa o indirectamente, esos objetos o situaciones puedan considerarse parte de las relaciones de preferencia. El valor es pues no una característica inherente a las cosas o las situaciones, no algo que existe como parte de la estructura ontológica del universo, de un modo completamente independiente de las personas y sus actividades. Antes bien, el valor se crea o se determina mediante la preferencia. Los valores son productos de nuestros afectos.

Concebir el valor como algo que depende de las relaciones afectivas es concebirlo como algo *subjetivo*. La concepción opuesta, es decir la concepción objetiva, no niega que los valores se relacionen con los afectos, pero sí niega que los valores sean el producto de tales relaciones. Concebir el valor como algo objetivo es sostener que el valor existe independientemente de los afectos de las personas conscientes y que proporciona una norma o regla que permite gobernar los afectos. El punto de vista subjetivista niega la existencia de semejante norma.

Aquí el lector podría preguntarse si la concepción teológica que considera el valor como algo que depende de la voluntad divina es objetiva o subjetiva.[7] Puede argumentarse que la dependencia de la voluntad de Dios es análoga a la dependencia del afecto consciente. Nosotros podemos replicar que en la mayor parte de las concepciones teológicas, todo depende de la voluntad divina, de modo tal que, desde el punto de vista de las

criaturas finitas y conscientes, el valor tiene el mismo carácter objetivo que otros aspectos del orden divino. Concebir el valor como una norma o regla ordenada por la divinidad que gobierna nuestro afectos, es considerarlo algo objetivo.

La diferencia entre las concepciones subjetiva y objetiva del valor no es una diferencia entre los objetos y las situaciones a los cuales se les asigna valor. Podría pensarse que el subjetivismo implica que sólo aquello que es subjetivo en sí mismo (sólo un estado de experiencia consciente) puede poseer valor intrínseco. Pero la preocupación del subjetivismo es el fundamento del valor y no su objeto. No hay ninguna restricción que penda sobre la naturaleza de aquellas situaciones que pueden ser objetos de preferencia y que, por lo tanto, pueden ser evaluadas. Pueden existir algunas personas que revelen y expresen preferencias sólo en relación con su propia experiencia, pero seguramente hay pocos que limitan de ese modo sus preferencias. No debemos confundir la idea de que sólo aquello que es preferido por su propio bien es o puede ser intrínsecamente bueno, con la idea de que sólo aquello que ha sido experimentado placenteramente puede ser intrínsecamente bueno.

Una concepción objetiva del valor es igualmente neutral respecto de la naturaleza de los objetos de valor intrínseco. Esa concepción es completamente compatible (aunque ciertamente no constituye una exigencia) con la idea de que sólo los portadores de valor intrínseco son estados de experiencia consciente. Uno podría suponer que la actividad contemplativa es la única, o más exactamente la suprema, situación intrínsecamente valorable y, mientras uno pueda afirmar que también es la más verdaderamente placentera y la más apropiadamente preferida, esos juicios referentes a su carácter afectivo podrían derivar de su valor, antes que determinarlo, lo cual, en esta perspectiva, dependería de la naturaleza cognitiva de la experiencia.

O podría considerarse una visión más mundana, es decir que la supervivencia de la conciencia individual es el único bien intrísecamente bueno. Un subjetivista podría apoyar esta posición porque supone que los seres conscientes, capaces de elegir, consideran en realidad la supervivencia de la conciencia como su único objetivo, por lo menos si la someten a una consideración plena. Un objetivista también podría apoyar esta posición, aun

cuando suponga que algunos seres conscientes no consideran la supervivencia como su único interés, pero insistirían en afirmar que sin embargo tal posición proporciona el objetivo adecuado para una elección racional. El subjetivista basaría el valor de la supervivencia consciente según el sitio real que le corresponda entre nuestras preocupaciones; el objetivista determinaría el auténtico lugar que ocupa la supervivencia consciente sobre la base de su valor. Pero ambos considerarían que la experiencia consciente proporciona el único valor intrínseco. Si ambos se equivocan (como sostendremos nosotros), su error no ofrece ningún fundamento para elegir entre las concepciones subjetivista y objetivista del valor.

El subjetivismo no debe confundirse con la idea de que los valores son arbitrarios. Como ya lo hemos hecho notar, si la utilidad ha de identificarse con el valor debe ser también la medida de la preferencia considerada. Los valores no son meros rótulos que pueden aplicarse al azar o caprichosamente a las situaciones, sino que antes bien son registros de las actitudes plenamente consideradas que sostenemos, según nuestras creencias, ante aquellas situaciones. Aunque la relación entre la creencia y la actitud no está en si misma abierta a la evaluación racional, no por ello es arbitraria. Hemos mencionado ya que estamos de acuerdo con la afirmación de Hume referente a que no es contrario a la razón preferir que se destruya el mundo a un raguño en el dedo. Una persona que después de haberlo considerado plenamente, sostiene tal preferencia, asignaría correctamente a la destrucción del mundo un valor mayor que el que podría asignarle a la destrucción de su dedo. Aunque, como también observamos antes, tendríamos fundamentos, independientes de la satisfacción de esa preferencia, para sostener que esa persona está loca, no tendríamos fundamentos para considerar que su preferencia es arbitraria. La locura no implica necesariamente ni una falla de la razón ni una falla de la reflexión. (La locura, en nuestra perspectiva, es primariamente un desorden de los afectos, pero ello no implica que los afectos sean irracionales, del mismo modo en que un desorden del estómago no implica que sea irracional.)

Tampoco debe confundirse el subjetivismo con la idea de que es imposible conocer los valores. La evaluación, como una

actividad de medición, es cognitiva. La preferencia, es decir lo que se mide, se puede conocer. Lo que niega el subjetivista es que haya un conocimiento del valor que no sea un conocimiento comunmente empírico, un conocimiento de un reino especial de lo valorable, al que sólo se puede llegar a través de cierta forma de intuición diferente de la experiencia de los sentidos. El conocimiento del valor corresponde únicamente a la esfera de los afectos. La evalución es cognitiva, pero no hay una única cognición "orientada al valor".

El valor es la norma de la elección. El objetivista considera el valor como una norma para la elección, porque también es una norma de la preferencia. Para él el hecho de relacionar la elección con la preferencia no muestra que la elección sea racional, salvo en la medida en que la preferencia en sí misma pueda relacionarse con el valor y ser su norma. El subjetivista considera el valor como una norma de la elección porque el valor es la medida de la preferencia considerada. Para él relacionar la elección con la preferencia, muestra que la elección es racional siempre que la preferencia haya sido considerada plenamente, sea coherente, que permanezca estable ante la reflexión y la experiencia y satisfaga las condiciones necesarias para definir una medida de intervalos. Para el objetivista la preferencia se interpone entre la elección y el valor; para el subjetivista, el valor se interpone entre la elección y la preferencia.

4.2 En la concepción que estamos defendiendo, el valor es una medida de la preferencia individual. El valor de una situación le indica a una determinada persona el lugar que ocupa en sus preferencias, se lo mide. Cada situación se caracteriza, como ya lo mencionamos en la primera sección de este capítulo, no por un valor único, sino por un conjunto de valores, uno por cada relación afectiva de la cual participa, o que se puede considerar que participa. La relación de los valores de las diferentes personas o entre los diferentes valores asignados a una situación particular, es una cuestión que corresponde a la investigación empírica. Puede pensarse que los valores divergentes reflejan creencias opuestas, pero es posible que no toda esta afirmación sea correcta. Sin embargo, también podría pensarse que los valores convergentes reflejan la aceptación no crítica de un individuo o de los miembros de un grupo de las ideas

de los demás, con la exclusión de la plena consideración individual. No tenemos ninguna razón valedera para suponer que las preferencias plenamente consideradas de las diferentes personas sobre las mismas situaciones den lugar a medidas equivalentes, ni siquiera suponiendo que las personas estén plenamente informadas, hayan reflexionado convenientemente, cuenten con la necesaria experiencia y se encuentren similarmente colocadas respecto de la situación dada.

Concebir el valor como algo dependiente de las relaciones afectivas propias de cada individuo es considerar que el valor es *relativo*. Por supuesto, hay otras concepciones relativistas del valor. Es común a todas las formas de relativismo puramente individualista la idea de que cada persona tiene su propio concepto del bien (y del mal) y que los bienes de las diferentes personas no son partes de un único bien superior común. Por otra parte, las formas no individualistas del relativismo pueden suponer que existe un bien común a una familia o a una clase o a una sociedad y negar que los bienes de las diferentes familias o clases o sociedades son partes de un bien común. Una concepción absoluta o universal del valor se opone pues a todas las formas del relativismo. Una concepción absolutista sostiene que los valores son los mismos para todas las personas o para todos los seres conscientes. Un absolutista puede suponer que los valores están determinados por los afectos, de modo tal que una situación es buena simplemente porque es preferida por cierta persona, pero entonces debe también sostener que la situación es buena no sólo en relación con esa persona o desde su punto de vista, sino también desde el punto de vista de todas las personas. El relativista individualista supone que el bien de uno y el bien de otro son absolutamente independientes y que no forman parte de un bien único universal. El absolutista insiste en afirmar que el bien de una determinada persona es simplemente esa parte del bien único que le concierne.[18] Si estuviéramos desarrollando una teoría completa del valor deberíamos examinar los matices tanto de la posición relativista como de la absolutista. Y nos encontraríamos no con dos opiniones opuestas, sino con un continuum de perspectivas. Pero nos quedaremos con la sencilla oposición que existe entre el relativismo puramente individualista y el absolutismo indiferenciado.

En ambas concepciones, la absolutista y la relativista, es posible distinguir qué es bueno *para* cierta persona de lo que es directamente bueno y suponer que una situación que es buena para cierto individuo puede considerarse buena desde el punto de vista de los demás. Una situación es buena para alguien si contribuye a su bienestar. En la perspectiva absoluta, si contribuye al bienestar de alguien es una fuente de valor positivo. Por lo tanto, es una fuente de valor positivo para todos. Una situación que es buena para un individuo, y quizá sólo buena para ese individuo, es por lo tanto buena desde el punto de vista de cualquiera, por ser buena para ese individuo. En la perspectiva relativa, el hecho de que una situación contribuya al bienestar de alguien es una fuente de valor positivo para otros si esa situación forma parte del bien de esas personas, como lo es en nuestra posición relativista particular si aquellas otras personas prefieren tal contribución. Una situación que es buena para un único individuo es pues buena desde el punto de vista de toda persona que prefiera o pueda preferir el aumento del bienestar de aquel individuo.

Quizá la formulación filosófica clásica de una concepción del valor tanto subjetiva como relativa ya fue ofrecida en el siglo XVII por Thomas Hobbes. En el *Leviathan* Hobbes dice:

> Pero sea cual fuere el objeto del Apetito o el Deseo de cualquier hombre; es decir, lo que por su parte él llama *Bueno*, y sea cual fuere el objeto de su Odio y Aversión, que él llama *Malo*...Porque estas palabras Bueno y Malo...siempre se emplean en relación con la persona que las usa. De modo que no hay nada simple ni absoluto, ni tampoco hay reglas comunes del Bien y del Mal que puedan tomarse de la naturaleza de los objetos mismos, tales reglas sólo pueden proceder de la Persona del hombre.[19]

Y luego vuelve Hobbes a decir:

> Lo *Bueno* y lo *Malo* son palabras que representan nuestros Apetitos y Aversiones, los cuales difieren según los distintos temperamentos, las distintas costumbres y las distintas doctrinas. Y los distintos hombres difieren no sólo en su Juicio, o sea en la sensación de qué es agradable y qué es desagradable al gusto, el olfato, el oído, el tacto y la vista, sino que también difieren en lo que consideran que agrada o desagrada a la Razón, en las acciones de la vida corriente. Más aun, el mismo hombre, en diferentes situaciones, difiere de sí mismo y en determinado momento aprecia, es decir, llama Bueno, a lo que en otro momento menosprecia y llama Malo.[20]

Hobbes vincula el subjetivismo con el relativismo. La idea de que el valor depende del apetito o la preferencia con la idea de que el valor es relativo a cada individuo. Por más natural que pueda parecer este vínculo, y por ello es que fácilmente se confunden las dos posiciones, cada una podría sostenerse sin la otra. En realidad, una interpretación común y seguramente errada del propio Hobbes niega que él mismo sea un subjetivista y sostiene que es un relativista.[21] Hobbes supone que el máximo bien desde el punto de vista de cualquier persona es su propia supervivencia. Como ya dijimos antes, ésta puede considerarse una interpretación objetiva; la supervivencia individual nos proporciona una norma para nuestros afectos y preferencias, de modo tal que todo aquel que busque su propia destrucción es un ser irracional. Incluso si, después de considerarlo plenamente, prefiere la muerte a la vida. Si ésta fuera la visión de Hobbes, sería difícil relacionarla con la posición expuesta en los párrafos citados en los cuales el valor está claramente vinculado con el apetito de un modo subjetivista.

En los enfoques expuestos por una de las teorías que más influencia tienen en la moral moderna, el utilitarismo, parecen implícitas tanto la posición subjetivista como la absolutista. John Stuart Mill sugiere una posición semejante en su intento de ofrecer una especie de prueba del principio de utilidad: subjetivista, al decir que "la única prueba que se puede presentar de que algo es deseable es que la gente realmente lo desee" y absolutista, al afirmar que "la felicidad de cada persona es un bien para esa persona, y por consiguiente, la felicidad general es un bien para el conjunto de todas las personas".[22] Pero en las propias declaraciones de Mill se advierte una evidente inconveniencia en esa unión entre el subjetivismo y el absolutismo, pues el hecho de que Mill pase de una premisa aparentemente relativista (que la felicidad de cada persona es un bien para esa persona) a una conclusión absolutista (que la felicidad general es un bien para todas las personas) ha sido generalmente tomado para ejemplificar la falacia de la concepción.

El utilitarismo se encuentra bajo la presión de apartarse de una concepción del valor al mismo tiempo subjetiva y absoluta. La forma más plausible de oponerse a esa presión es aparentemente aceptar una concepción universalista de la racionalidad y sostener que puesto que la racionalidad se identifica con la maximización del valor y la racionalidad es universal, lo que se maximiza, es decir, el valor, debe ser pues similarmente universal —el mismo desde el punto de vista que se mire—. Si, con todo, en la concepción económica de la racionalidad, el utilitarismo continúa siendo verdadero hasta en sus raíces, tanto el absolutismo como el subjetivismo se desvanecen. Por un lado, el valor puede concebirse como algo relativo, pero se presenta una forma especial de valor, el valor moral, que constituye la medida de aquellas preferencias consideradas sostenidas desde un punto de vista especialmente limitado para asegurar la imparcialidad. Por el otro lado puede concebirse el valor como algo objetivo, como la medida de una característica inherente al estado de la experiencia —el goce— que suministra una pauta o norma para la preferencia. Este no es el lugar indicado para iniciar una discusión sobre esas posiciones, de modo que afirmaremos meramente, pero de forma dogmática, que cien años de esfuerzos cada vez más rebuscados para evitar la falacia de Mill no hicieron avanzar la causa del utilitarismo ni un solo centímetro. Pero, por supuesto, debemos prestar una atención más seria especialmente a la segunda de las dos formas mencionadas antes de defender el utilitarismo y lo haremos al continuar la exposición de nuestra teoría.

Si bien es cierto que existe una presión contra una concepción del valor simultáneamente subjetiva y absoluta, existe también una presión contra una concepción al mismo tiempo objetiva y relativa. La idea de que cada individuo es un miembro de una especie natural y que cada especie tiene su propia perfección característica, completamente diferente de las otras especies naturales, no está hoy ampliamente difundida en la ética secular, sin embargo es de gran importancia histórica. Pero esta visión raramente —si alguna vez lo hizo— se adhirió a una concepción relativa del valor. Porque siempre se supuso que la objetividad de cada perfección característica, el papel que desempeña como norma o pauta según la cual puede juzgarse a

cada miembro individual de la especie, dependía de considerar cada perfección como manifestación (apropiada para esas circunstancias particulares) de un bien universal único. Los bienes aparentemente relativos de varias especies son en realidad facetas del bien absoluto. Por consiguiente, las demandas de objetividad conducen a una concepción aparentemente relativa del valor, dentro de un molde absolutista.

Por lo tanto, trataremos a la concepción relativa del valor como la contraparte natural de una concepción subjetiva, aun cuando ambas sean lógicamente independientes. Pero, el relativismo puede parecer una doctrina absurda. Una manzana McIntosh es algo bueno de comer (en realidad, yo mismo estoy disfrutando de una mientras escribo estas palabras). Si otra persona negara esto podría estar equivocada. Sin embargo esa persona puede no ser un amante de las manzanas, o su gusto por las manzanas puede llegar hasta la dulce Golden Delicious y no hasta la ácida McIntosh. Al decir que una McIntosh es algo bueno de comer, la palabra bueno no está "utilizada únicamente en relación con la persona que la emplea". En realidad uno podría equivocarse al declarar que una manzana McIntosh es una buena manzana si el único fundamento en el que se basa es su propio gusto. De ahí que aparentemente no se trate de una evaluación relativa. Sin embargo es subjetiva, la bondad de las manzanas McIntosh es independiente del gusto de cualquier individuo en particular, pero no del gusto en general. El valor de una manzana McIntosh es una medida de preferencia, pero no de una preferencia común.

Por más plausible que pueda parecer, este argumento contra el relativismo pone de manifiesto una confusión (presente en el pensamiento de Hobbes y bastante común en las formulaciones de una posición relativista) que puede llevar a una objeción directamente errada. Hobbes se equivoca al suponer que palabras tales como "bueno" se utilizan en relación con la persona que la emplea; en realidad nuestros términos típicamente evaluativos suponen un punto de vista común que puede o no ser compartido por quien evalúa. De modo que una buena manzana es aquella que se prefiere común o usualmente para comer. Pero, aun estando comprometida con este punto de vista por el lenguaje de la evaluación, una persona no está comprome-

tida con él ni como una medida de sus propias preferencias, ni como una pauta para sus propias elecciones. Al decir que una McIntosh es una buena manzana para comer, una persona no está implicando que sea una amante de las manzanas, ni que disfrute de las McIntosh más que de la mayoría de las otras manzanas, ni que elegiría para comer una manzana McIntosh. Una manzana McIntosh es buena en relación con el punto de vista supuesto en el uso de la palabra bueno, pero ese valor no nos brinda una norma para la elección desde cualquier punto de vista.

La utilidad de una situación es el valor que tiene para una persona como un resultado posible, como una alternativa, por lo tanto, de otras situaciones posibles ciertas. Comenzando por estos valores de la utilidad, subjetivos y relativos como son, podemos proseguir en direcciones completamente diferentes. Podemos generalizar desde las utilidades que miden las preferencias de una persona particular, referentes a las diferentes situaciones, hasta el tipo de valores que sostiene esa persona. De esos valores personales no podemos inferir directamente las preferencias particulares, sino más bien los tipos de preferencia que la persona ha de manifestar en diferentes situaciones. Podemos estimar esos valores personales de los modos analizados en las secciones anteriores de este capítulo y determinar hasta qué punto son valores considerados y coherentes.

Pero, también podemos generalizar desde las utilidades que miden las preferencias de los individuos particulares hasta la serie de valores manifestados por las diferentes situaciones de las que pueden formar parte un objeto particular o un tipo de objeto. Vemos que en algunos casos las preferencias de las diferentes personas proceden de puntos de vista compartidos o similares; esto queda puesto de manifiesto por órdenes paralelos de sus utilidades. El lenguaje común del valor muestra esos órdenes paralelos y podemos hablar de manzanas buenas para comer porque encontramos similitudes en los gustos de aquellos que comen manzanas.

Sin embargo, es posible establecer un lenguaje común del valor de un modo más sistemático, sin necesidad de observar gustos paralelos. Pues podemos considerar que los objetos, o por lo menos algunos de los objetos que forman parte de los resul-

tados directamente valuados, se relacionan mutuamente en un mercado. Hay mercancías, objetos de intercambio entre las personas. Así concebido, cada objeto llega a poseer un valor fijo en el intercambio, valor que lo relaciona con todas las demás mercancías y que es común para todas las personas cuyas relaciones de intercambio están englobadas en ese mercado.

El punto de vista en el que hay que situarse para evaluar, inducido por el mercado, no sólo es reconocido por cada una de las personas que hace su evaluación, sino que, en realidad, es compartido. Las preferencias paralelas entre los amantes de las manzanas determinan un punto de vista común, desde el cual una manzana McIntosh es evaluada como una buena manzana para comer; cualquiera puede reconocer esta evaluación sin compartir la preferencia por las McIntosh. Pero si una McIntosh tiene un mayor valor de mercado que una Golden Delicious, si cuesta 25 centavos y una Golden sólo 20, sea cual fuere el gusto de la persona que evalúa, ésta preferiría una McIntosh, puesto que puede intercambiarla por una Golden Delicious y recibir además 5 centavos. Preferir una Golden Delicious a una McIntosh, en el contexto del intercambio, equivaldría a preferir 20 centavos a 25. Pero los valores comunes y los puntos de vista comunes del mercado se basan en las utilidades de los individuos que son estrictamente relativos. Al comprender cómo los valores absolutos o universales del mercado surgen de un conjunto dado de valores individuales relativos y de una tecnología de producción dada, en realidad comprendemos la falacia en la que incurre el relativismo del valor que ignora las normas comunes que están en la base de la evaluación corriente. Puesto que la teoría del valor en el intercambio es un lugar común de la economía elemental, no es el caso exponerla aquí, pero sí es la teoría a la que debe remitirse el lector si se siente tentado por las objeciones más comunes que suelen hacérsele a la concepción relativista del valor.

4.3 La defensa del subjetivismo es la tarea primaria que debe afrontar quien identifica el valor con la medida de la preferencia considerada. La defensa del relativismo es esencialmente un asunto de prolija indagación de la relación que mantiene el individuo con las normas comunes. Pero el conflicto entre el enfoque subjetivo y el enfoque objetivo del valor se

encuentra en el corazón mismo de la filosofía moral. Ningún desacuerdo obstaculiza más seriamente el desarrollo de una adecuada teoría moral. No podemos esperar resolver este conflicto de un modo que a todos les parezca convincente o siquiera plausible. En el caso del subjetivismo sin embargo nos parece apremiante hacerlo y trataremos de explicar nuestra razones —razones que, como lo hicimos notar al principio, le deben mucho al trabajo del Gilbert Harman y de John Mackie—.[23]

El primer problema que se nos plantea es formular la posición opuesta, es decir, la objetivista. Corremos pues el inevitable peligro de crear un hombre de paja. Particularmente debemos cuidarnos de atribuirle al objetivista compromisos ontológicos, y luego defender el subjetivismo mediante el rechazo de esa ontología, pues el objetivista considera esos compromisos innecesarios o evitables. Digamos que el objetivista sostiene que cualquier explicación adecuada de nuestra experiencia o de nuestro ambiente debe referirse al valor o a lo valorable, considerado como algo independiente de los seres sensibles y de sus afectos. Desde nuestro punto de vista, esto compromete al objetivista con una ontología del valor, pero no nos basaremos en ese punto para desarrollar nuestro argumento.

Ante todo debemos aclarar qué está implícito en la independencia del valor que se le atribuye al objetivismo. El objetivista puede afirmar que existe un vínculo necesario entre los seres sensibles y el valor y sostener que el valor nos proporciona la norma de nuestros afectos. El objeto adecuado de preferencia es, y necesariamente debe serlo, el bien. Pero, desde el punto de vista del objetivista, ese vínculo no se encuentra dentro del afecto consciente, sino que es un lazo que ata el afecto a algo más que en sí mismo proporciona el fundamento de tal atadura. Esa otra punta del lazo (y no debemos imponer al objetivista una explicación particular de qué es eso) debe ser independiente de los seres conscientes y de sus afectos, aun cuando los vincule entre sí. En nuestro análisis designaremos a esa otra punta supuesta del lazo con la expresión *valor objetivo*.

¿Cómo hemos de decidir si cualquier explicación adecuada de nuestra experiencia o de nuestro ambiente debe referirse al valor objetivo? En nuestra perspectiva, la respuesta es que

consideremos si, en *la mejor explicación* que podemos dar de nuestras acciones y elecciones existe alguna referencia al valor objetivo.[24] No son éstos el lugar ni el momento indicados para embarcarnos en una teoría de la explicación, de modo que solamente diremos que el único esquema serio que permite dar una explicación de la acción humana es, como ya lo mencionamos, el siguiente: la elección maximiza la satisfacción de la preferencia, según una creencia dada. El valor objetivo no desempeña ningún papel en esta explicación. La mejor explicación que podemos dar de nuestras observaciones es que hay objetos físicos con propiedades que, gracias a nuestro aparato sensorial, nos provocan esas observaciones. Por consiguiente, aceptamos las propiedades físicas como parte de cualquier explicación adecuada de nuestra experiencia y nuestro ambiente. Pero la mejor explicación que podemos dar de nuestras elecciones no es que existen objetos con valores que, gracias a nuestro aparato afectivo, nos hacen decidir esas elecciones. Por lo tanto, rechazamos los valores objetivos como parte de cualquier explicación adecuada de nuestra experiencia y nuestro ambiente. La diferencia entre las notorias propiedades simples de G. E. Moore, amarillo y bueno, es que desde el punto de vista de la explicación lo bueno es una quinta rueda.[25] El valor objetivo, como el flogisto, es una parte innecesaria de nuestro aparato explicativo, y como tal debe rasurarse de la faz del universo, mediante la navaja de Ockham.

John Mackie ha sostenido que el mejor ejemplo contra el valor objetivo resulta ser su rareza ontológica.[26] Esta no se adapta cómodamente a nuestra explicación de la aprehensión o del conocimiento, no se relaciona felizmente con el "hecho". El objetivista puede lamentarse de que Mackie le imponga exigencias ontológicas y epistemológicas que no son parte de su posición. Pero el valor objetivo puede ser, si no ya raro, por lo menos único en cuanto a la clase de fuerza o poder que es capaz de ejercer. El valor objetivo debe ejercer una forma de atracción como el objeto mismo de preferencia que es muy diferente de cualquier otra atracción que observamos o experimentamos. (Sin embargo, ¿es así? ¿No podría parecerse la atracción del valor objetivo a la atracción que ejercen ciertas personas? "Desde el momento en que la mujer entró en la habitación, todos

los hombres sólo tuvieron ojos para ella. Uno no podía decir que su apariencia, o sus ademanes o su vestido, fueran en sí mismos llamativos, pero nadie podía ignorar a aquella mujer.") Sin embargo, no deberíamos considerar la rareza como algo que descalifique el valor objetivo, dadas las partículas y las fuerzas a las cuales se refieren nuestras mejores explicaciones físicas. Lo raro del valor objetivo es que no desempeñe ninguna función explicativa y eso es lo que lo descalifica.

Mackie sugiere que "si la doctrina teológica requerida pudiese defenderse", podría ser cierta una forma objetiva del valor.[27] Esto parece correcto, puesto que la defensa de tal doctrina teológica constituiría o incluiría una defensa de su función explicativa. Y si relacionamos el valor con la voluntad divina, en semejante doctrina el valor objetivo podría pues cumplir también una función explicativa. No podemos negar que en principio la mejor explicación para nuestras acciones y nuestras elecciones puede exigir una referencia al valor objetivo; en realidad nuestros principales antepasados intelectuales (los griegos clásicos y los pensadores europeos de la Edad Media) dieron explicaciones precisamente con esa exigencia. Pero nuestras explicaciones no son las de aquellos pensadores. Cuando Laplace condenó a Dios a la redundancia explicativa diciendo que "je n'avais pas besoin de cet hypothèse", también estaba condenando implícitamente el valor objetivo.[28]

Hay una objeción evidente a lo que hemos dicho en cuanto a que el valor objetivo no cumple una función explicativa. Las atribuciones de valor frecuentemente forman parte de las razones ofrecidas para explicar nuestras elecciones y acciones. "¿Por qué envió usted a su hijo a Harvard?" "Porque es una buena universidad." Esta no es una explicación muy esclarecedora, pero, por lo menos, es el comienzo de una explicación, mientras que "Porque prefiero a Harvard" no constituye ninguna explicación.

Parte del papel explicativo de las atribuciones de valor puede comprenderse si uno se basa en nuestro análisis de un lenguaje común del valor que expusimos en la subsección anterior. Relacionamos nuestras elecciones y acciones con aquello que comunmente se prefiere o con aquello que se prefiere en el intercambio. En la medida en que los criterios para tal

preferencia sean conocidos por todos es imposible pues inferir de las atribuciones de valor las características particulares; un buena cuchilla es aguda y filosa. Una concepción subjetiva y relativa del valor da cuenta rápidamente de la función explicativa de tales atribuciones de valor.

Admitimos que se considera que las atribuciones de valor poseen un rol explicativo adicional. Pues la creencia en el valor objetivo se sostiene amplia aunque confusamente y aquellos que creen en el valor objetivo apelan a él cuando quieren explicar su conducta, del mismo modo en que aquellos que creen en Dios apelan a El para explicar su conducta. Pero, aunque en la explicación que suministramos de ciertas elecciones y acciones debemos referirnos a la creencia en el valor objetivo, no por ello estamos claramente comprometidos con tal creencia. El hecho de que no podamos comprender lo que las personas religiosas hacen sin referirnos a su creencia en Dios, no nos compromete con tal creencia. (Que no podamos explicar lo que hacen ciertos geómetras, sin referirnos a su creencia en la posibilidad de la cuadratura del círculo, no nos compromete con tal creencia.) Entre las creencias que forman parte de la explicación de nuestras elecciones hay algunas que, si fueran verdaderas, deberían necesariamente referirse al valor objetivo. Lo que negamos es que tales creencias sean verdaderas.

Nuestra explicación del valor debe incluir por consiguiente lo que Mackie llama "teoría del error".[29] Suponemos que las personas objetivan sus preferencias y así llegan a considerar sus actitudes y afectos subjetivos como propiedades que caracterizan los objetos de sus preferencias. Si supusiéramos que la concepción correcta del valor podría descubrirse mediante un análisis del lenguaje común, sin duda llegaríamos a una concepción objetiva (o quizás a una concepción con elementos objetivos y subjetivos entremezclados en irremediable confusión). Pero si, en cambio, suponemos que la concepción correcta del valor sólo puede descubrirse recurriendo a la mejor explicación de lo que se supone que afecta el valor, pondríamos pues al descubierto el error presente en los enfoques corrientes y estableceríamos una concepción subjetiva. Seguramente, el hecho de inferir la mejor explicación no ha de mostrar que no existe un valor objetivo, del mismo modo en que tampoco ha de demostrar que no hay hadas

en el fondo del jardín. Nos parece una buena ideas colocar el valor objetivo en el mismo plano que a las hadas.

Consideramos pues que el valor es una medida de la preferencia individual —subjetivo, porque es una medida de la preferencia y relativo, porque es una medida de la preferencia individual—. Lo que es bueno, en esencia lo es porque es algo preferido y es bueno desde el punto de vista de aquellos, y sólo de aquellos, que lo prefieren. Hemos tratado de incluir algunas calificaciones y exactitudes a este enfoque con el fin de resultar convincentes, pero, lo repetimos, un desarrollo completo y una defensa acabada exigirían un estudio más profundo y no sólo este breve esbozo. Nuestra intención es demostrar la posibilidad y las características de una moral racional, considerando que el valor es en sí mismo subjetivo y relativo.

Notas

1 Shakespeare, *Hamlet,* II, ii.255-7.

2 Hume, *Treatise,* ii,iii,iii, pag.415.

3 Sobre la distinción entre la racionalidad paramétrica y la racionalidad estratégica, véase J. Elster, *Ulysses and the Sirens: Studies in rationality and irrationality,* Cambridge, 1979, págs. 18-19,117-23.

4 Véase J. C. Harsanyi, "Advances in Understanding Rational Behavior", en *Essays on Ethics, Social Behavior, and Scientific Explanation,* Dordrecht, 1976, pág. 94.

5 Véase Harsanyi, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations* Cambridge, 1977, pág.9.

6 Hume, *Treatise,* ii,iii,iii, pág. 416.

7 Hay diferencias significativas en el uso de los términos "subjetivo" y "objetivo" y "relativo" y "absoluto". Trataremos de aclarar el sentido que les damos en la sección 4.1-2 *infra.*

8 Véase R. D. Luce y H. Raiffa, *Games and Decisions,* Nueva York, 1957, págs. 50-51, y D. M. Winch, *Analytical Welfare Economics* Harmondsworth, Middx., 1971, pág. 25.

9 "Suponemos que los individuos tienden a maximizar la utilidad y definimos la utilidad como aquello que el individuo tiende a maximizar", Winch, pág. 25.

10 Lo que nos importa son los tipos ideales, antes que las posiciones particulares expuestas en la bibliografía. Puede encontrarse un análisis útil basado en el interés, que rechaza la concepción

objetiva del valor pero reduce el interés a la satisfacción, en L. W. Sumner, *Abortion and Moral Theory,* Princeton, 1981, págs. 162-186. Obsérvese que nuestra propia referencia al interés hecha en el capítulo 1 pretendía defender la preferencia y la utilidad y no el concepto específico discutido aquí. Puede encontrarse un análisis sobre la influencia de la prudencia en T. Nagel, *The Possibility of Altruism,* Oxford, 1970, págs. 27-76.

11 Sumner afirma, "si algo aumenta mi bienestar, debe producirme, directa o indirectamente, inmediatamente o a largo plazo, estados del espíritu que yo considero agradables o evitarme estados del espíritu que considero desagradables", pág.184. De modo que si el valor es la medida del bienestar individual, o de la felicidad o de la satisfacción, el mandato conceptual establece una concepción hedonista del valor.

12 "La perspectiva correcta de la propia vida es la del *momento actual*", B. Williams, "Persons, character and morality", en *Moral Luck,* Cambridge, 1981, pág. 13. Véase, además del artículo de Williams, D. Parfit, "Later Selves and Moral Principles", en A. Montefiore (comp.), *Philosophy and Personal Relations,* Londres, 1973, págs. 137-169. Véase también el análisis de la posición que sostiene que "el concepto del hombre racional debería ampliarse de modo tal que incluyera consideraciones de índole temporal", en Elster, págs. 65-77.

13 Véase Luce y Raiffa, capítulo 2 y Harsanyi, *Rational Behavior,* capítulo 3.

14 Véase Harsanyi, *Rational Behavior,* pág. 9.

15 Luce y Raiffa, *op. cit.,* pág. 27.

16 Elster continúa el pasaje citado en la nota 14, capítulo 1 del siguiente modo: "Aún faltan plantear muchos problemas referentes a su objetivo y su fuerza de resolución". Luego Elster ofrece "una lista de tales problemas". Y no es una lista breve.

17 Mi colega Shelly Kagan justamente se hacía ésta y otras preguntas destinadas a discutir y clarificar muchos otros puntos.

18 Véase G. E. Moore, *Principia Ethica,* Cambridge, 1903, págs. 97-102, sobre una defensa de esa forma de absolutismo.

19 Hobbes, *Leviathan,* cap. 6, pág. 24.

20 *Ibíd.,* cap. 15, pág. 79.

21 Véase la introducción de B. Gert a Hobbes, *Man and Citizen,* Nueva York, 1972, págs. 13-16.

22 John Stuart Mill, *Utilitarianism,* Londres, 1863, capítulo 4, párrafo 3.

23 Véase G. Harman, *The Nature of Morality: An Introduction to Ethics,* Nueva York, 1977, capítulo 1; también J. L. Mackie, *Ethics: Inventing Right and Wrong,* Harmondsworth, Middx., 1977, capítulo 1.

24 Nuestro enfoque ya fue sugerido por Harman, *The Nature of Morality*, págs. 6-9. La frase "la mejor explicación" es suya, véase "The Inference to the Best Explanation", *Philosophical Review* 74, 1965, págs. 88-95.

25 Véase Moore *op. cit.*, págs. 7-10 sobre su comparación de amarillo y bueno.

26 Véase Mackie *op. cit.*, págs. 38-42.

27 *Ibid.*, pág. 48.

28 Véase E.T. Bell, *Men of Mathematics*, 2 vols., Harmondsworth, Middx., 1953, vol.1, pág. 198.

29 Véase Mackie, *op. cit.*, págs. 35, 48-49.

III

Estrategia: razón y equilibrio

1.1 Jane desea intensamente asistir a la fiesta de Ann. Pero desea aún más evitar a Brian que puede estar allí. Brian desea intensamente evitar la fiesta de Ann. Pero desea aún más encontrarse con Jane. Si espera que Brian asista a la fiesta de Ann, Jane permanecerá en su casa. Si Brian espera que Jane se quede en su casa, también él permanecerá en la suya. Si Jane espera que Brian permanezca en su casa, irá a la fiesta. Si Brian espera que Jane asista, también lo hará él. Si Jane...pero aquí llegamos adonde empezamos. Desde puntos de vista opuestos, Jane y Brian afrontan un problema de interacción. Si cada uno conoce no sólo sus propias preferencias sino también las del otro, ¿qué expectativa debería tener? y ¿cuál de las posibles acciones debería elegir?

En este capítulo nos concentraremos en el tema de la racionalidad estratégica: la racionalidad en la interacción. En nuestro análisis de la racionalidad paramétrica, en la cual el actor considera que sus circunstancias son fijas y por lo tanto su elección es el único elemento variable, comprobamos que la definición de una medida de intervalos de la preferencia sobre los posibles resultados, nos permite reducir el problema de la elección racional a un ejercicio de maximizar la utilidad esperada. La interacción se opone a un tratamiento tan directo. Los actores racionales deben determinar sus elecciones no en circunstancias fijas, sino atendiendo a las expectativas recíprocas sobre aquellas mismas elecciones. Tanto las elecciones como las expectativas deben basarse en las creencias que tienen los actores referentes a las acciones que pueden realizar, los resultados posibles de tales acciones y las utilidades que pueden brindar esos posibles resultados. En el caso ideal —y por improbable que pueda ser un ideal acabado, nos suministra un punto de apoyo para teorizar sobre la elección racional— éstas son cuestiones de conocimiento común. Cada uno sabe que los demás las conocen. Pero entonces, el razonamiento que cada

persona hace partiendo de esos datos sobre sus propias expectativas y elecciones debe ser accesible a cualquier otra persona. En efecto, la elección individual debe surgir del razonamiento común. Cada actor debe considerar que su elección estratégica constituye una respuesta a las elecciones de los demás actores y recibe a su vez una respuesta constituida por las elecciones de éstos.

Este caso ideal será la base sobre la que desarrollaremos nuestro análisis de la racionalidad estratégica.[1] Las complicaciones que provocan una racionalidad imperfecta y un conocimiento incompleto ocultan la estructura lógica de la interacción, y es esta estructura lo que sirve como fundamento de una teoría moral. El único punto en el que nos apartaremos del caso ideal será decir que ese caso es subjetivo. Esto es, examinaremos la interacción desde el punto de vista de un individuo que supone que prevalecen las condiciones ideales. Ese individuo es una persona racional, supone que sus pares también son racionales y cree que la racionalidad de todos los actores es una cuestión de conocimiento común. Nuestro individuo supone que cada uno de los actores tiene una perfecta información sobre las acciones posibles, sobre los resultados posibles y sobre todas las preferencias que existen acerca de tales resultados, y también sabe que es de conocimiento común que cada uno de los actores posee esa información perfecta. El individuo en cuestión introduce una medida de intervalo para evaluar las preferencias de cada persona, define una función de utilidad para cada una y supone que todos los demás hacen lo propio. Por consiguiente, cree que su razonamiento referente a la interacción constituye una réplica al razonamiento de cada uno de los demás y a su vez recibe la réplica del razonamiento de éstos. Nuestro análisis de la racionalidad estratégica parte de ese punto de vista ideal, pero para nuestra discusión es irrelevante el hecho de que el punto de vista mismo se base en creencias y suposiciones correctas.

A fin de precisar aun más nuestro análisis, formulamos tres condiciones que debe cumplir la elección estratégicamente racional:

> *A*: La elección de cada persona debe ser una respuesta racional a las elecciones que esa persona espera que hagan los demás.

> *B*: Cada persona debe esperar que la elección de cada una de las demás satisfaga la condición *A*.
> *C*: Cada persona debe creer que su elección y sus expectativas se reflejan en las expectativas de cada una de las demás.

La condición *A* relaciona la racionalidad del actor con el marco de interacción; y le exige a la persona ser estratégicamente racional. La condición *B* expone de manera explícita el supuesto de que todos los que participan en la interacción son seres estratégicamente racionales y que esa racionalidad es una cuestión de conocimiento común. La condición *C* manifiesta explícitamente el supuesto de que cada persona considera la situación como si tuviera un conocimiento completo de los fundamentos de la elección, como si ese conocimiento fuera compartido por todos y como si todos supieran que comparten tal conocimiento. Cada condición es necesaria, pero no afirmamos aquí que cada una por su parte sea suficiente para la interacción racional.

Al formular las tres condiciones hemos dejado sin definir la frase "respuesta racional". Si identificamos la racionalidad con la maximización de la utilidad individual podríamos suponer que bastaría sustituir sencillamente en la condición *A* la palabra "racional" por la expresión "maximización de la utilidad". Hagámoslo y relacionemos las condiciones así entendidas con nuestro ejemplo inicial.

Jane tiene dos elecciones —ir a la fiesta de Ann o quedarse en su casa—. Supongamos que Jane decide asistir. Entonces, por la condición *A* debe esperar que Brian decida permanecer en su casa. Así, por la condición *B*, Jane debe esperar que Brian espere que ella se quede en su casa. Pero de ese modo su elección no se refleja en lo que Jane cree que son las expectativas de Brian, de modo que se viola la condición *C*. Supongamos pues que Jane decide quedarse en su casa. Entonces, por la condición *A*, debe esperar que Brian elija asistir a la fiesta de Ann. Y por la condición *B*, Jane debe esperar que Brian espere que ella asista. Pero entonces, la elección de Jane no queda reflejada en lo que ella cree que son las expectativas de Brian, de modo que nuevamente se viola la condición *C*. Ninguna de las posibles elecciones de Jane satisface nuestras tres condiciones. Tampoco lo hace ninguna de las elecciones de Brian. Entonces, ¿cuál sería

la elección racional posible para Jane y para Brian? ¿Cómo podemos formular un principio de racionalidad estratégica que satisfaga las condiciones *A*, *B* y *C*?

1.2 Antes de poder contestar a estas preguntas debemos presentar algunas convenciones terminológicas y gráficas.[2] Aunque tales convenciones entorpezcan un poco la lectura, son necesarias para comprender nuestra argumentación, no sólo en éste sino también en los capítulos siguientes.

Una acción es el objeto de una elección. Un resultado, como ya lo definimos antes, es el producto de una acción y de un conjunto de circunstancias determinadas. Cuando se realiza una elección en una situación de certeza, cada acción está correlacionada con un único conjunto de circunstancias y por lo tanto con un único resultado. Cuando la elección se realiza en una situación de riesgo o de incertidumbre, cada acción está correlacionada con una distribución de la probabilidad referente a varios conjuntos de circunstancias, de modo tal que puede compararse con una lotería cuyos premios son los resultados posibles. Continuamente hemos supuesto que desde el punto de vista de un actor individual, cada acción posible tiene un resultado distinto para cada conjunto de circunstancias al cual el actor le atribuye una probabilidad diferente de cero.

Ahora modificamos esta concepción de lo que es un resultado. En la interacción un resultado procede no de la elección de un actor único, sino de las elecciones de varios actores. Por consiguiente, inicialmente modificamos nuestra definición de un resultado de modo tal que decimos que un resultado es el producto de varias acciones, una correspondiente a cada persona implicada en la interacción, y que también es el producto de un conjunto de circunstancias determinadas. Pero esta nueva definición es innecesariamente molesta. La modificamos aun más si suprimimos la última cláusula, es decir, si decimos sencillamente que un resultado es el producto de varias acciones correspondientes cada una a cada persona implicada en la interacción.

¿Cómo justificamos pues el hecho de haber quitado toda referencia a las circunstancias? Por supuesto, un actor particular puede atribuir una probabilidad distinta de cero a cada uno de los varios conjuntos de circunstancias determinadas

correlacionadas con una interacción. Entonces, ese individuo supone que cada conjunto cuyos elementos son acciones (cada una elegida por cada persona que participa de la interacción) puede tener varios resultados posibles. Pero para cada actor tal conjunto determina una única utilidad igual a la utilidad que él espera para el conjunto considerado como una lotería que afecta esos posibles resultados. Nosotros consideramos esa utilidad única como si fuera la utilidad de un resultado único y definimos ese resultado como el producto de los elementos del conjunto de acciones. De este modo, eliminamos la referencia a las circunstancias.

Habrá de notarse que en la interacción las elecciones particulares o las acciones particulares forman parte de dos relaciones completamente diferentes. Por un lado, si consideramos una acción por cada una de las personas que participan de la interacción, tenemos pues un conjunto de acciones simultáneamente posibles y hemos definido el resultado como el producto de los elementos de un conjunto semejante. Por otro lado, si consideramos todas las acciones posibles para cualquiera de las personas que participan de la interacción, tenemos pues un conjunto de acciones que se excluyen recíprocamente: las elecciones posibles de un actor dado. Un resultado es pues el producto que se obtiene tomando una elección de cada uno de esos conjuntos. Por consiguiente, el conjunto de resultados que pueden darse en cualquier interacción es el producto cartesiano del conjunto de elecciones posibles que tienen los actores.

Podemos representar convenientemente cualquier interacción que incluya sólo a dos personas mediante un esquema en el cual representamos horizontalmente en una hilera las elecciones o acciones de una persona y, encolumnadas verticalmente, las elecciones o acciones de la otra persona. Cada intersección de una hilera horizontal y una columna vertical representa pues un resultado, y en esa intersección podemos mostrar la utilidad que el resultado tiene para cada una de las personas. Por convención las utilidades de la hilera se presentan primero. Cuando suponemos únicamente una medida ordinal de preferencia debemos emplear un orden jerárquico, es decir que debemos colocar primero el resultado más preferido. Pero cuando suponemos una medida de intervalos debemos emplear

los números cardinales para representar las utilidades. Debe recordarse que esos números sólo tienen una significación de intervalos. Tanto el punto cero como la unidad de medición se seleccionan arbitrariamente. (Comparémoslo con la medición de la temperatura. La diferencia entre 30 grados centígrados y 50 grados centígrados es el doble que la diferencia que hay entre los 20 y los 30 grados centígrados. Pero 40 grados no es el doble de calor que 20 grados.) Ademas, las comparaciones a intervalos sólo pueden hacerse entre las utilidades de un mismo individuo; la medida no proporciona una base para la comparación interpersonal.

Podemos ilustrar la representación esquemática de la interacción utilizando nuestro ejemplo inicial. Un orden posible de las preferencias de Jane y de Brian podría ser el siguiente:

		Brian			
		asistir		quedarse en casa	
Jane	asistir	4ta	1ra	1ra	4ta
	quedarse en casa	2da	3ra	3ra	2da

Partiendo de los datos que tenemos, las preferencias primera y cuarta de Jane deberían ser claras. Jane prefiere permanecer en su casa si Brian asiste a la fiesta de Ann antes que permanecer en casa si Brian no asiste a la fiesta, pues en el último caso lamentaría no haber ido a la fiesta. De modo que tenemos sus preferencias segunda y tercera. Las preferencias primera y segunda de Brian deberían ser claras partiendo de los datos que tenemos. Brian también prefiere ir a la fiesta si Jane se queda en su casa antes que permanecer en su casa si Jane asiste a la fiesta, aun cuando la primera situación le proporcione una velada menos placentera, porque en el segundo caso lamentaría mucho no haber asistido a la fiesta. De modo que tenemos las preferencias tercera y cuarta.

El esquema nos permite ver que si Brian asiste a la fiesta, la respuesta de Jane que maximiza su utilidad es permanecer en casa. Si Jane se queda en su casa, la respuesta de Brian que maximiza su utilidad es permanecer también en su casa. Si Brian se queda en su casa, la respuesta de Jane que maximiza su utilidad es ir a la fiesta. Y si Jane va, la respuesta de Brian

que maximiza su utilidad es también ir a la fiesta. Por consiguiente si la acción que realiza una persona es una respuesta a la acción de la otra que maximiza su utilidad, la acción que realiza la segunda persona no constituye una respuesta a la acción de la primera que maximice su utilidad. Esto es suficiente para asegurar que no se satisfacen simultáneamente las condiciones *A*, *B* y *C*.

Como un primer paso para mostrar que sin embargo la elección racional es posible en situaciones como ésta, debemos ampliar el alcance de la elección, de modo tal que cada individuo pueda elegir no solamente una acción, sino también una lotería sobre las posibles acciones. En el capítulo anterior identificamos la acción realizada en situación de riesgo o incertidumbre como una lotería cuyos premios eran los resultados posibles. Ahora presentamos la idea de una elección que funciona como una lotería cuyos premios son las acciones posibles. Pero no deben confundirse estos dos tipos de lotería.

La ampliación del alcance de la elección que permite incluir loterías cuyos premios son las acciones, podría ser inútil en las situaciones paramétricas, pero es esencial para la posibilidad misma de que se haga una elección racional en las situaciones estratégicas. Si tomamos la elección de una lotería la acción que se realiza verdaderamente queda determinada por un artificio fortuito programado con la distribución elegida de los premios. Jane podría elegir una lotería que le dé al hecho de asistir a la fiesta una probabilidad de 1/4 y al hecho de permanecer en casa una probabilidad de 3/4 seleccionando un naipe al azar de un mazo sin comodines, por ejemplo que decida ir a la fiesta si y sólo si el naipe extraído es una espada. Suponemos aquí la disponibilidad universal de un artificio fortuito programable para asegurar la distribución de probabilidades exigida por cualquier lotería, y suponemos también que el artificio puede utilizarse sin costo alguno. Una vez más suponemos estas situaciones ideales a fin de simplificar nuestro análisis.

Llamamos una *estrategia* a una lotería sobre las acciones posibles. Una *estrategia pura* atribuye la probabilidad 1 a una acción y la probabilidad 0 a las demás acciones. Es una lotería con un único premio y ese premio siempre aparece. Una *estrategia mixta* atribuye una probabilidad distinta de cero a más de

una acción y, por supuesto, la suma de las probabilidades asignadas es igual a 1; es una lotería que tiene por premios varias acciones.

El hecho de ampliar el alcance de la elección de las acciones a las estrategias no extiende la cantidad de resultados posibles. Sea cual fuere la estrategia que elija cada actor, éste debe realizar una de sus acciones posibles y el producto de las acciones realizadas por cada actor es uno de los resultados posibles. Sin embargo, ese resultado no queda determinado con certeza por la elección de las estrategias salvo que se elijan únicamente estrategias puras. Cada estrategia es una lotería sobre las acciones, de modo que cada conjunto de estrategias (cada una correspondiente a cada persona que participa en la interacción) determina una lotería sobre los resultados. Definimos un *resultado esperado* de tal lotería como el producto de las loterías o las estrategias elegidas por cada persona. Por consiguiente, el conjunto de resultados esperados es el producto cartesiano de los conjuntos de estrategias de los actores. De ahora en adelante, utilizaremos el termino "resultado" para indicar el "resultado esperado", salvo en los casos en que el contexto lo excluya claramente.

Ahora podemos definir el concepto clave indispensable para explicar la racionalidad estratégica: el concepto de *equilibrio*. Un resultado (esperado) está en equilibrio o es un resultado en equilibrio si y sólo si es el producto de estrategias, estrategias que maximizan cada una la utilidad esperada de la persona que la eligió atendiendo a las estrategias elegidas por las demás personas. Para simplificar esto podríamos decir que un resultado está en equilibrio si y sólo si es el producto de estrategias que maximizan mutuamente las utilidades. Por lo tanto, si establecemos las condiciones *A*, *B* y *C* e interpretamos que una respuesta racional es una respuesta que maximiza la utilidad, cada persona debe esperar que el resultado de su elección esté en equilibrio.

La demostración de esta exigencia es directa. Supongamos que hay n personas y tomemos a cualquier persona, a quien podemos designar como la persona 1, que elige una estrategia s_1. Por la condición *A*, s_1 maximiza la utilidad de esa persona dadas las estrategias $s_2,...s_n$ que son las que esa persona espera

que los demás elijan. Por la condición B, s_2 debe maximizar la utilidad de la persona 2 dadas las estrategias $s'_2, s'_3,...,s'_n$ que son las que la persona 1 espera que la persona 2 espere. Y s_3 debe maximizar la utilidad de la persona 3 dadas las estrategias $s"_1$, $s"_2$, $s"_4,...s"_n$ que son las que la persona 1 espera que la persona 3 espere y así sucesivamente. Por la condición C, la elección y las expectativas de la persona 1 deben quedar reflejadas en sus creencias referentes a las expectativas de los demás, de modo tal que $s_1=s'_1=s"_1$... y $s_2=s"_2$=..., y $s_3=s'_3$=..., y así sucesivamente. De ahí que cada estrategia $s_1,s_2,..., s_n$ debe ser maximizadora de la utilidad dadas las otras, y así el resultado esperado de la elección s_1 debe estar en equilibrio.

Importante como es este argumento en cuanto a relacionar la racionalidad estratégica con el equilibrio, debemos tener presentes sus supuestos y sus limitaciones. La comprobación depende tanto de considerar a las respuestas racionales como maximizadoras de la utilidad, como de suponer que cada persona se forma determinadas expectativas referentes a las elecciones de los demás. Y la comprobación no muestra que los resultados de las estrategias racionalmente elegidas deben estar en realidad en equilibrio. Esto podría solamente darse si cada persona conociera las utilidades de las demás, de modo tal que todas las expectativas fueran materia del conocimiento común. Las estrategias racionalmente elegidas pueden basarse en expectativas erradas, de modo tal que el verdadero resultado puede no estar necesariamente en equilibrio. Pero entonces la falla está en las expectativas y no en la elección.

Volvamos una vez más a nuestro ejemplo inicial. Podemos observar que ninguno de los resultados que proporcionan las estrategias puras está en equilibrio. A fin de introducir las estrategias mixtas y de atribuir utilidades esperadas a sus resultados, necesitamos una medida de intervalos de la preferencia. Esto exige contar con datos adicionales; debemos relacionar las preferencias interpuestas en los órdenes de Jane y de Brian con las loterías que obran sobre sus preferencias extremas. Supongamos pues que Jane es indiferente entre su segunda preferencia (permanecer en su casa si Brian asiste a la fiesta) y una lotería que le asigne 2/3 de probabilidad a su primera preferencia (ir a la fiesta si Brian se queda en su casa) y 1/3 de

probabilidad a su cuarta preferencia. Empleando el método que vimos en la sección 3.2 del capítulo 2, podemos asignarles utilidades de 1, 2/3, 1/3 y 0 a los cuatro resultados según el orden de las preferencias de Jane. Supongamos que Brian es indiferente entre su segunda preferencia (que ambos se queden en sus casas) y una lotería que le asigne 1/2 probabilidad a su primera preferencia (que ambos asistan a la fiesta) y 1/2 probabilidad a su cuarta preferencia (que se quede en su casa y Jane vaya a la fiesta). Y supongamos también que Brian es indiferente entre su tercera preferencia (ir a la fiesta y que Jane se quede en su casa) y una lotería que le asigne 1/6 de probabilidad a su primera preferencia y 5/6 de probabilidad a su cuarta preferencia. Podemos pues asignarles utilidades de 1, 1/2, 1/6 y 0 a los cuatro resultados según el orden de preferencias de Brian. Podemos representar esta situación, que suponemos es la que sostienen tanto Jane como Brian, del siguiente modo:

		Brian			
		asistir		quedarse en casa	
Jane	asistir	0	1	1	0
	quedarse en casa	2/3	1/6	1/3	1/2

Supongamos que Jane elige una estrategia mixta que le asigna una probabilidad de 1/4 a ir a la fiesta y una probabilidad de 3/4 a permanecer en casa. Entonces, si Brian va a la fiesta su utilidad esperada es [(1/4 x 1) + (3/4 x 1/6)] = 3/8. Si Brian permanece en su casa su utilidad esperada es [(1/4 x 0) + (3/4 x 1/2)] = 3/8. Cada acción le proporciona a Brian la misma utilidad esperada. Elija lo que eligiere Brian, su utilidad esperada será 3/8.

Supongamos que Brian elige una estrategia mixta que le asigna una probabilidad de 1/2 al hecho de asistir a la fiesta y una probabilidad de 1/2 al hecho de quedarse en casa. El lector puede verificar que elija lo que eligiere Jane, su utilidad esperada será 1/2. De ello se sigue que el efecto que resulta de la elección de estas dos estrategias mixtas está en equilibrio. Cada una como respuesta a la otra es maximizadora de la utilidad. Por supuesto, cada una de esas respuestas no es la única que maximiza la utilidad, pero nuestra definición no exige

que se cumpla eso. (Si Jane eligió una estrategia, cualquier respuesta que dé Brian debería maximizar la utilidad y si Brian ya hizo una elección, cualquier respuesta que dé Jane debería maximizar la utilidad.)

Al extender sus elección a fin de que éstas incluyan loterías cuyos premios son las acciones, hemos mostrado que Jane y Brian pueden llegar a obtener un resultado que esté en equilibrio. Por supuesto, podemos verificar si se satisfacen las condiciones *A*, *B* y *C* considerando que una respuesta racional maximiza la utilidad. Cada elección es una respuesta a la elección esperada de la otra persona que maximiza la utilidad. Cada uno espera que la elección del otro sea una respuesta maximizadora de la utilidad a la expectativa que ese otro tiene de la elección del primero. Y cada uno cree que su elección y sus expectativas están reflejadas en las expectativas del otro.

Nuestro ejemplo ilustra una posibilidad general. J. F. Nash demostró que en cualquier situación que cuente con un número finito de personas (lo cual no llega a ser un supuesto restrictivo), cada una de ellas con cantidad finita de estrategias puras y con preferencias que satisfacen las condiciones necesarias para establecer una medida de intervalos, debe haber por lo menos un resultado (esperado) en equilibrio.[3] Debe haber por lo menos un conjunto de estrategias que maximicen mutuamente las utilidades. Por lo tanto, en cualquier interacción , cada persona puede relacionar su elección y sus expectativas sobre las elecciones y las expectativas de los demás con un conjunto de estrategias que maximicen mutuamente las utilidades. Siempre es posible satisfacer las condiciones *A*, *B* y *C* si consideramos que una respuesta racional es una respuesta que maximiza las utilidades. Aparentemente Nash demostró que en principio la elección estratégicamente racional siempre es posible.

El principio de elección racional que surge de este análisis es el de que en la interacción cada persona debería relacionar su elección de la estrategia y sus expectativas de las elecciones que han de hacer los demás con un conjunto de estrategias que maximicen recíprocamente las utilidades. Este principio no siempre constituye una guía práctica para la acción; quizá no haya un modo factible de determinar un conjunto estratégico adecuado. En una situación compleja, en la que hay muchas

personas, que a su vez cuentan con muchas estrategias puras, el costo de determinar el equilibrio puede resultar mucho mayor que cualquier beneficio que pueda alcanzarse si se logra conocer tal equilibrio. Pero, dejando de lado esa cuestión práctica, pueden hacerse varias objeciones teóricas al principio propuesto. Esas objeciones están vinculadas con las premisas de las cuales deriva tal principio y se refieren a la supuesta conexión entre una respuesta racional y una respuesta que maximiza la utilidad o a la suposición de que cada persona puede formarse una determinada expectativa sobre la estrategia elegida por cada una de las demás personas.

2.1 Hemos interpretado que la condición *A* exige que cada persona elija una estrategia que maximice su utilidad esperada, atendiendo a sus expectativas referentes a las elecciones de los demás. Pero esa persona puede tener varias estrategias semejantes. Entonces la condición *A* permite elegir cualquiera de ellas y considerar a cada una como una respuesta racional. Pero esto es incompatible con el principio que exige que la preferencia quede puesta de manifiesto mediante la elección racional. Si un individuo tiene varias estrategias (y cada una es maximizadora de la utilidad, según sus expectativas) debería ser indiferente a cualquiera de los resultados posibles. En la conducta, esa indiferencia se expresa mediante una elección entre las estrategias basada en idénticas probabilidades. Si la persona eligiera una estrategia en lugar de otra (excepto si se tratara del resultado de una lotería) esa actitud revelaría una preferencia del individuo por el resultado de la estrategia elegida, lo cual le exigiría asignarle a ese resultado una utilidad esperada mayor que la que le atribuye a los resultados de las estrategias rechazadas. Y esto sería incompatible con el supuesto de que cada estrategia es igualmente maximizadora de la utilidad.[4]

Si aceptamos esta objeción, debemos pues reformular la condición *A* y decir que la elección de cada persona debe ser *la* respuesta racional a las elecciones que esa persona espera que los demás hagan; en ese caso definiríamos *la* respuesta racional como la respuesta determinada por una lotería que le asigna idénticas probabilidades a todas las respuestas que satisfagan las exigencias de racionalidad del otro. Puesto que identificamos la racionalidad con la maximizacion de la utilidad, podemos

decir que la elección de cada persona debe ser *la respuesta centroide maximizadora de la utilidad* a las elecciones esperadas de los demás. Y así debemos reformular el principio propuesto para la elección estratégicamente racional y decir que cada persona debe relacionar su elección de una estrategia y sus expectativas de las elecciones de los demás con un conjunto de estrategias centroides mutuamente maximizadoras de la utilidad.

Pero este principio no puede cumplirse en muchas situaciones en las que no hay estrategias centroides mutuamente maximizadoras de la utilidad. Recordemos el ejemplo analizado en la primera sección de este capítulo. Vimos que se llega al equilibrio si Jane elige una estrategia mixta que le asigna una probabilidad de 1/4 al hecho de asistir a la fiesta de Ann y una probabilidad de 3/4 al hecho de permanecer en su casa, y Brian elige una estrategia mixta que le asigna una probabilidad de 1/2 al hecho de asistir a la fiesta y una probabilidad de 1/2 al hecho de quedarse en casa. En esta situación no hay otro equilibrio posible. Si Jane aumentara su probabilidad de asistir a la fiesta de Ann, la respuesta maximizadora de la utilidad de Brian sería ir a la fiesta, y la respuesta maximizadora de la utilidad de Jane a esto sería quedarse en su casa y así repetiríamos el ciclo como nos ocurrió al comienzo del capítulo. Si Jane quisiera disminuir su probabilidad de asistir a la fiesta, la respuesta maximizadora de la utilidad de Brian a eso sería permanecer en su casa y la respuesta maximizadora de la utilidad de Jane a eso sería asistir a la fiesta con lo cual caemos nuevamente en el mismo circulo. Si Brian decidiera cambiar su estrategia, el resultado sería el mismo.

Pero, como ya hemos observado, si Brian elige una estrategia que le asigna las mismas probabilidades al hecho de asistir a la fiesta que al hecho de quedarse en casa, cualquier respuesta que decida dar Jane ha de ser maximizadora de su utilidad. Por lo tanto, la respuesta centroide maximizadora de la utilidad de Jane es su estrategia mixta que le asigna las mismas probabilidades a ir a la fiesta que a quedarse en casa. Pero entonces la estrategia de Brian no constituye una respuesta maximizadora de la utilidad a la nueva estrategia de Jane; pues Brian debería asistir a la fiesta. En esta situación no hay estrategias centroides mutuamente maximizadoras de la utilidad.

Hemos dicho que un resultado está en equilibrio si y sólo si ese resultado es el producto de estrategias mutuamente maximizadoras de la utilidad. Y decimos que un resultado está en *fuerte* equilibrio si y sólo si cada estrategia es la única respuesta maximizadora de la utilidad que el actor puede dar a las estrategias del otro. De lo contrario, el resultado está en un equilibrio *débil*. Si un resultado está en equilibrio fuerte, éste es pues trivialmente el producto de un conjunto de estrategias centroides mutuamente maximizadoras de la utilidad. Pero si un resultado está en equilibrio débil, no necesita ser, y generalmente no lo es, el producto de un conjunto de estrategias centroides mutuamente maximizadoras de la utilidad. Puesto que muchas situaciones, tales como la de nuestro ejemplo, carecen de un equilibrio fuerte, consideramos inaceptable la reformulación propuesta de nuestras condiciones. Pero no queremos dejar sin definir la elección racional en situaciones que sólo tienen resultados en equilibrio débil.[5]

Nosotros no negamos la exigencia de que la preferencia quede puesta de manifiesto en la elección racional, pero hacemos notar que esa exigencia debe aplicarse con mucho cuidado a la racionalidad estratégica. La elección revela una preferencia determinada por una creencia. En la interacción, la creencia incluye la expectativa. Una persona puede estar genuinamente indiferente ante varias estrategias según sus expectativas. Pero esas expectativas pueden reflejar la elección de una y sólo una de las estrategias que le son indiferentes. La persona debe pues elegir esa estrategia a fin de sustentar su expectativa. Si Jane espera que Brian elija entre asistir a la fiesta y permanecer en su casa con idénticas probabilidades, equivale a decir que Jane es indiferente ante las estrategias posibles. Pero, por la condición *B*, Jane debe esperar que la elección que haga Brian maximice la utilidad del joven, según las expectativas de Brian y la estrategia elegida por Jane. Y, por la condición *C*, Jane debe esperar que Brian espere la estrategia que elige Jane. El hecho de que ir a la fiesta y quedarse en casa tengan las mismas probabilidades maximiza la utilidad de Brian si y sólo si Brian espera que Jane elija la estrategia que le asigna una probabilidad de 1/4 al hecho de que Jane asista a la fiesta y una probabilidad de 3/4 al hecho de que ella se quede en su casa. De

modo que entre las estrategias que le resultan indiferentes a Jane, sólo una sustenta la expectativa sobre la que se basa la indiferencia. La exigencia de que la preferencia quede puesta de manifiesto en la elección racional, sólo ha de satisfacerse de un modo conveniente al carácter estratégico de la interacción.

2.2 Nash prueba que en cualquier situación hay por lo menos un conjunto de estrategias mutuamente maximizadoras de la utilidad o, como consideraremos de ahora en adelante más apropiado decir, por lo menos un conjunto en equilibrio de estrategias. Pero existe la posibilidad de que haya más de un conjunto con esa característica. En realidad deberíamos trazar brevemente una representación esquemática de una situación en la cual cada conjunto que contenga una estrategia por cada persona implicada en la interacción es un conjunto en equilibrio, aunque la configuración de preferencias que daría lugar a tal representación pueda ser poco común. El principio de elección racional propuesto al final de la primera sección de este capítulo exige solamente que cada persona relacione su elección y sus expectativas sobre las elecciones de los demás con un conjunto en equilibrio. Pero, si hay varios conjuntos semejantes, ¿cuál es el que debe elegir la persona? ¿Con qué conjunto debería relacionar sus expectativas referentes a las elecciones de los otros, de modo tal que la propia elección constituya una respuesta racional a las de los demás? La persona podría no formarse expectativas sobre las elecciones de los demás sencillamente a voluntad, y entonces, ¿cómo podría fundamentar sus expectativas?

Si sólo hay un conjunto de estrategias en equilibrio, puesto que sólo las expectativas relacionadas con ese conjunto satisfacen las condiciones *A*, *B* y *C* (suponiendo que una respuesta racional debe ser maximizadora de la utilidad) esas expectativas tienen fundamentos suficientes. El primer paso que hay que dar para fundamentar las expectativas, cuando hay varios conjuntos en equilibrio, es considerar un caso especial en el cual sólo un conjunto proporciona un resultado en equilibrio *no dominado*. Decimos que un resultado en equilibrio es dominado si cada uno de los actores implicados prefiere cualquier otro resultado en equilibrio. Puesto que ningún resultado se domina a sí mismo y que la dominación es una relación transitiva, un

resultado en equilibrio no puede pues estar dominado sólo por otros resultados en equilibrio dominados; si un resultado es dominado significa que debe haber un resultado en equilibrio no dominado que lo domina. Por consiguiente, todos los actores que participan de la interacción deben preferir un único resultado en equilibrio no dominado a cualquier otro resultado en equilibrio. Ese resultado suministra un foco natural a las expectativas de las personas que identifican la racionalidad con la maximización de la utilidad. Cada uno de los actores elegiría la estrategia que correspondiera al conjunto que proporciona el único resultado en equilibrio no dominado, y esperaría que los demás hicieran lo propio. Aunque esto no es algo que exijan las condiciones *A*, *B* y *C*, parece constituir una extensión no problemática de la concepción de la elección racional que expresan tales condiciones.

El segundo paso que hay que dar para fundamentar las expectativas cuando existen varios conjuntos en equilibrio es considerar el caso en el cual hay un solo conjunto que proporciona un resultado en equilibrio no dominado que todos prefieren a cualquier otro resultado en equilibrio. Si sólo hay un resultado en equilibrio no dominado y sólo uno, como ya lo mostramos, cada actor debe pues preferirlo a cualquier otro resultado en equilibrio. Consideremos ahora el caso en el que puede haber varios resultados en equilibrio no dominados, pero sólo uno preferido por todos a cualquier otro resultado. Si llamamos a ese resultado *Y*, y lo comparamos con cualquier otro resultado en equilibrio no dominado que llamamos *X*, tendremos que algunas personas son indiferentes entre *Y* y *X*, algunas prefieren *Y* a *X* y ninguna prefiere *X* a *Y*. Una vez más parecería que semejante resultado proporciona un foco natural a las expectativas. Cada actor elegiría la estrategia correspondiente al conjunto que proporcione el único resultado en equilibrio preferido por todos a cualquier otro resultado en equilibrio, y esperaría que los demás hagan lo propio. Pero esto demuestra ser una extensión algo más problemática de la concepción de la elección racional expresada por las condiciones *A*, *B* y *C* que podrían haberse pensado al principio.

Consideremos la situación que se presenta en este esquema:

		Norman	
		lo hace	no lo hace
Noreen	lo hace	1,1	0,1
	no lo hace	1,0	0,0

Todo resultado esperado de esta situación está en equilibrio. Cualquier resultado esperado en el cual Noreen o Norman hagan lo que fuere que estuvieran decidiendo hacer es un resultado en equilibrio no dominado. Pero sólo el resultado que se da si ambos lo hacen está en equilibrio no dominado y es preferido por ambos a cualquier otro resultado en equilibrio. Si aceptamos el argumento del párrafo precedente, este resultado es pues un foco natural para las expectativas de Noreen y de Norman, de modo que ambos deberían hacer lo que fuere que estaban decidiendo hacer.

Pero parecería que ninguno de los dos tiene razón para hacerlo. En esta situación, ninguna de las dos posibilidades puede afectar la utilidad que obtiene Noreen mediante su elección. El único efecto que puede ejercer la elección de cada persona es determinar la utilidad del otro y esto no les da una razón para elegir a las personas que identifican la racionalidad con la maximización de la utilidad individual. En esta situación la exigencia de que la elección ponga de manifiesto la preferencia parecería determinar que cada uno elija una lotería que le asigne idéntica probabilidad a hacerlo que a no hacerlo; en otras palabras, que cada actor elija la respuesta que maximice su utilidad centroide. Esto le brindaría a cada uno una utilidad esperada no de 1 sino de 0,5.

En este ejemplo salen a la superficie cuestiones profundas de nuestro enfoque de la racionalidad y, en realidad, de la moral. La utilidad es una medida de la preferencia individual. Puesto que no se le impone ninguna restricción a la satisfacción de la preferencia, la preocupación exclusiva que cada persona manifieste por su propia utilidad no implica ningún egoísmo material. Pero este ejemplo sugiere que la estructura de algunas interacciones conduce a un egoísmo formal. Tanto Noreen como Norman pueden maximizar la utilidad del otro sin que tal actitud resulte un costo personal, pero parecería que ninguno tiene una razón para hacerlo.

En próximos capítulos desarrollaremos un enfoque de la

cooperación racional que puede aplicarse a esta situación, a fin de demostrar que sería racional para Noreen y para Norman hacer la elección que maximice la utilidad del otro. Pero no podemos hacerlo ahora. De modo que dejaremos de lado esta situación y volveremos al problema general de fundamentar las expectativas cuando tenemos varios conjuntos de estrategias en equilibrio.

Consideremos la situación representada por este esquema:

			Valerie
		b_1	b_2
Victor	a_1	1, 1/3	0, 0
	a_2	1/3, 2/3	2/3, 1

A simple vista queda claro que (a_1 x b_1) y (a_2 x b_2) son resultados en equilibrio firme. Hay un tercer resultado que está en equilibrio débil. Si Victor elige su estrategia mixta que le asigna idénticas probabilidades a a_1 y a a_2, Valerie recibe una utilidad esperada de 1/2 sea cual fuere la estrategia que elija. De manera similar, si Valerie elige su estrategia mixta que le asigna idénticas probabilidades a b_1 y a b_2, Victor recibe pues una utilidad esperada de 1/2 sea cual fuere la estrategia que elija. Por lo tanto, esta estrategias proporcionan un resultado en equilibrio, [(1/2a_1 + 1/2a_2) x (1/2b_1 + 1/2b_2)].

Es evidente que la preferencia de Victor se inclina por (a_1 x b_1) y que la de Valerie lo hace por (a_2 x b_2). Pero si Victor eligiera su primera acción (a_1) con la optimista expectativa de que Valerie elija b_1, y Valerie eligiera su segunda acción con la expectativa optimista de que Victor elija a_2, el resultado real sería pues el desastroso (a_1 x b_2). Es decir, el peor de todos.

Supongamos primero que ninguno de los dos actores es capaz de formarse una determinada expectativa sobre la elección de estrategia que hará el otro. Ninguno de los dos podría pues determinar su respuesta maximizadora de la utilidad. ¿Hay pues otras bases para una respuesta que pueda considerarse racional? *Faute de mieux,* cada actor puede tratar de protegerse eligiendo la estrategia capaz de maximizar la utilidad mínima que ha de recibir, sea cual fuere la elección que haga el otro. Victor evitaría a_1, pues tal estrategia le proporcionaría 0 si Valerie eligiera b_2 pero puede hacer algo mejor que elegir a_2,

que le proporcionaría un 1/3 si Valerie eligiera a_1, es decir Victor puede elegir la estrategia mixta (1/4a_1 + 3/4a_2) lo que le brinda una utilidad esperada de 1/2 sea cual fuera la estrategia elegida por Valerie. A su vez Valerie evitaría b_2 y no podría hacer nada mejor que elegir b_1, que le proporcionaría 1/3 si Victor eligiera a_1, mientras que cualquier estrategia mixta le brindaría una utilidad menor que a_1. De este modo Victor puede garantizarse una utilidad esperada de 1/2 y Valerie, a su vez, puede garantizarse una utilidad esperada de 1/3; llamamos a éstas las utilidades maximínimas.

El recurso del maximínimo parece ser una decisión desesperada. Pero en realidad ofrece un modo de generar las expectativas que buscan Victor y Valerie. Victor compara tres perspectivas: la utilidad que le brinda su equilibrio más favorecido (1), la utilidad que le brinda el equilibrio más favorecido de Valerie (2/3) y la utilidad que le brinda su maximínimo (1/2). Si Victor aceptara el equilibrio más favorecido de Valerie, el costo de tomar como base su propio equilibrio más favorecido sería una utilidad de 1/3; si en cambio Victor alzara los brazos desesperanzado y optara por su maximínimo, el costo sería de 1/2. El promedio de esos costos, 2/3, puede considerarse una medida de su relativa renuencia a conciliar con Valerie.

Esa medida de la renuencia de Victor a conciliar con Valerie no depende de aquellos rasgos de la medida numérica de su utilidad —el cero punto y la unidad— que son arbitrarios. Cualquier medida de intervalo de las preferencias de Victor debe proporcionar el mismo grado de renuencia. Y esto nos suministra una base para poder hacer comparaciones interpersonales. Podemos preguntar si la renuencia de Victor a conciliar con Valerie es mayor o menor que la renuencia de Valerie a conciliar con él.

También Valerie compara tres perspectivas: la utilidad que le brinda su equilibrio más favorecido (1), la utilidad que le brinda el equilibrio más favorecido de Victor (1/3) y la utilidad que le brinda su maximínimo (1/3). Puesto que estos dos últimos son iguales, resulta evidente que a Valerie le cuesta lo mismo aceptar el equilibrio más favorecido de Victor que elegir su propio maximínimo. De modo que su renuencia relativa a conciliar con Victor —el promedio de los dos costos— es 1.

Podemos referirnos aquí al principio que deriva del trabajo de Frederik Zeuthen, quien afirma que una persona para quien el promedio entre el costo de la concesión y el costo del desacuerdo insuperable es menor que 1, racionalmente debe conciliar con el otro.[6] Esa persona está más dispuesta o es menos renuente a conciliar. En nuestro ejemplo esa persona es Victor. Así que Victor y Valerie comprueban que cada uno tiene una base para determinar sus expectativas referentes a la elección del otro. Victor espera que Valerie busque con firmeza su equilibrio más favorecido, es decir que espera que Valerie elija b_2; es decir que Victor elige a_2. Valerie espera que Victor ceda, de modo que espera que él elija a_2; entonces ella elige b_2. El tercer resultado, el que está en equilibrio débil no tiene cabida en nuestra argumentación puesto que ninguno de los dos actores lo prefiere a (a_2 x b_2).

Al ilustrar el empleo de las utilidades maximínimas y el principio de Zeuthen referente a la concesión racional destinado a determinar las expectativas sobre las elecciones en situaciones que cuentan con más de un conjunto de estrategias en equilibrio, sólo intentamos presentar un tema complejo y controvertido que no podemos continuar analizando aquí. El hecho de que se den varios resultados en equilibrio exige que los actores *coordinen* un equilibrio. El lector particularmente interesado en profundizar el estudio de los diferentes enfoques con que se ha tratado el problema de la coordinación puede consultar las obras de John Harsanyi y de Thomas Schelling y mi propio artículo titulado "Coordination".[7] Aquí sólo apuntamos a establecer que la exigencia de que la elección de una persona debe relacionarse con un conjunto de estrategias en equilibrio no queda restringida por el hecho de confiar en determinadas expectativas sobre situaciones que tienen un solo resultado en equilibrio y aun aquellas en las que hay un solo resultado en equilibrio no dominado.

Pero antes de abandonar este tema queremos señalar que los recursos presentados al analizar el problema de la coordinación que afrontan Victor y Valerie ha de desempeñar un papel central en el desarrollo de nuestra teoría moral. Más adelante comprobaremos que la función electiva que nos suministra un principio de negociación racional e imparcial depende de una

medida que se define como un promedio de las diferencias entre las utilidades, parecido pues al promedio de los costos empleado aquí. Y utilizaremos una forma ampliada del procedimiento que emplea Zeuthen para la concesión racional. Llegados a este punto examinaremos la racionalidad del principio de Zeuthen, un aspecto que, como el lector habrá advertido, hemos ignorado aquí.

3.1 El equilibrio es una propiedad de racionalidad de los resultados considerados como productos de las estrategias. Pero también es posible considerar los resultados como conjuntos de utilidades. Cada persona cuya estrategia es una de las contribuciones destinadas a producir un resultado recibe un beneficio que nosotros llamamos sus utilidades, como uno de los rendimientos del resultado. Y podemos definir una propiedad de racionalidad de los resultados atendiendo a sus rendimientos. Esa propiedad, que llamamos *optimidad,* se opone a la idea de identificar la "respuesta que maximiza la utilidad" con la "respuesta racional" que cumple las condiciones *A*, *B* y *C*.[8]

Un resultado (esperado) es óptimo (o más concretamente, Pareto-óptimo) si y sólo si no existe ningún resultado posible que le brinde a determinada persona una utilidad mayor y que no le brinde a nadie una utilidad menor. Alternativamente, un resultado es óptimo si y sólo si cualquier otro resultado que le brinde a alguna persona una utilidad mayor también le proporciona a alguna otra persona una utilidad menor. Podemos caracterizar la propiedad de optimidad de un modo que revela un significativo paralelo con el equilibrio. Si un resultado está en equilibrio, cada una de las estrategias que contribuyeron a él le brinda al actor que la eligió su utilidad máxima, suponiendo que no se haya alterado ninguna otra estrategia. Si un resultado es óptimo, cada uno de los beneficios que rinde proporciona a quien lo recibe su utilidad máxima, suponiendo que no decrezca ningún otro rendimiento. Las dos propiedades no son completamente paralelas: un resultado que le brinda a cada persona su utilidad máxima en el caso de que no se altere ningún otro rendimiento, puede no ser óptimo si todos los rendimientos pueden incrementarse simultáneamente. Pero cada propiedad representa un conjunto de máximos restringidos; los elementos de un conjunto de estrategias en equilibrio y los elementos de un

conjunto óptimo de rendimientos son ambos maximizadores de la utilidad mutua.

Si un resultado es óptimo no existe pues ninguno alternativo que sea unánimemente preferido. Pero si un resultado no es óptimo, debe haber entonces uno alternativo que le brinde mayor utilidad a cierta persona y que no le brinde a nadie una utilidad menor; y quizás haya uno que le brinde mayores utilidades a todos los actores implicados en la interacción. Un resultado no óptimo es ineficiente desde el punto de vista de su rendimiento utilitarista para aquellos que participan de la interacción. Por consiguiente, la optimidad es una propiedad atractiva para el resultado de la interacción.

Pero no debemos exagerar sus pretensiones. A cada resultado no óptimo le corresponde una alternativa óptima que beneficia a alguno de los participantes sin perjudicar a los demás. Pero no toda alternativa óptima de un resultado dado proporciona únicamente mayores beneficios. Volvamos a nuestros amigos Jane y Brian. Cada resultado que pueda darse si Jane decide ir a la fiesta —sea lo que fuere que elija Brian— es óptimo. Y todos ellos son resultados óptimos. Sin embargo, si consideramos el resultado de que ambos se queden en sus casas, comprobamos que cada actor espera una mayor utilidad si Jane asiste a la fiesta y si Brian elige una estrategia mixta que le asigna al hecho de asistir a la fiesta una probabilidad que oscila entre 1/2 y 2/3. (Una probabilidad de más de 1/2 le proporciona a Brian una utilidad esperada de más de 1/2, lo cual es la utilidad que le toca si ambos permanecen en su casa; una probabilidad de menos de 2/3 le proporciona a Jane una utilidad de más de 1/3, que es la misma utilidad que le toca si ambos permanecen es sus casas.) Pero si Brian elige cualquier otra estrategia, el resultado, aunque óptimo, les proporciona tanto a Jane como a Brian una utilidad esperada menor de la que recibiría Jane si ambos se quedaran en sus casas. Si Brian decide asistir a la fiesta (y supone que Jane también ha de asistir) el resultado es pues óptimo —cualquier alternativa le proporcionaría una utilidad menor—, pero es el peor que puede dársele a Jane.

El único resultado en equilibrio que existe en esta situación le brinda a Jane una utilidad esperada de 1/2 y a Brian una utilidad esperada de 3/8, como lo señalamos en la primera

sección. Es fácil advertir que éste no es un resultado óptimo. Si Jane decide asistir a la fiesta y Brian elige una estrategia mixta que le asigna al hecho de asistir a la fiesta una probabilidad que oscila entre 3/8 y 1/2, cada uno espera pues una utilidad mayor de la que podría esperar de un resultado en equilibrio. La posibilidad de que el/los resultado/s en equilibrio y el/los resultado/s óptimos de una situación particular se excluyan mutuamente ha de ser una parte fundamental de nuestra argumentación sobre la moral racional.

Como lo muestra este ejemplo, muchas situaciones presentan una multiplicidad de resultados óptimos. Aunque no siempre: Noreen y Norman sólo cuentan con una posibilidad óptima; cada uno debe hacer lo que sea que están decidiendo hacer. Aunque no tratemos de demostrar eso aquí, Valerie y Victor cuentan con dos conjuntos de resultados óptimos: los que se dan si Valerie elige b_1 y Victor elige una estrategia mixta que le asigna a a_1 una probabilidad mayor que 1/2 y el único resultado $(a_2 \text{ x } b_2)$. En esta situación, tanto el equilibrio preferido de Victor $(a_1 \text{ x } b_1)$ como el equilibrio preferido de Valerie $(a_2 \text{ x } b_2)$ son óptimos.

Dada la usual multiplicidad de los resultados óptimos y como además puede darse que un resultado minimice la utilidad de algún individuo y sin embargo sea óptimo, no pretendemos establecer que lograr la optimidad sea una condición suficiente para hacer una elección racional en la interacción. Pero sí podemos decir que es una condición necesaria. De modo que podría afirmarse que dadas las condiciones *A*, *B* y *C* una "respuesta racional" equivale a una "respuesta óptima". Una respuesta optimizadora es una estrategia cuyo resultado es óptimo, dadas las estrategias esperadas de las demás personas. Sólo ciertas expectativas son compatibles con una respuesta optimizadora; sólo si espera que Jane decida asistir a la fiesta Brian puede elegir una estrategia cuyo resultado sea óptimo. Pero toda situación, o por lo menos toda situación que implique solamente a un número finito de personas, cada una de las cuales sólo tiene un número finito de estrategias puras, debe dar por lo menos un resultado óptimo. Por lo tanto siempre es posible encontrar una estrategia y un conjunto de expectativas que satisfagan las condiciones *A*, *B* y *C* si se identifica la racionalidad con la optimización.

Ya hemos observado que la optimidad y el equilibrio no siempre son compatibles. Por lo tanto, si suponemos que la elección estratégicamente racional es generalmente posible, no podemos exigir que una respuesta sea al mismo tiempo maximizadora de la utilidad y optimizadora. Asistir a la fiesta es la respuesta optimizadora de Jane, sea cual fuere la elección que haga Brian, pero no puede combinarse con ninguna estrategia de Brian para establecer un par de respuestas que sean maximizadoras de la utilidad mutua. ¿Cómo podemos pues conciliar, o a falta de conciliación, decidir entre las demandas de maximización de la utilidad y de optimización? Este es el problema central que se le plantea a cualquier teoría de la racionalidad estratégica. Como mostraremos luego, la teoría moral es esencialmente la teoría de las limitaciones optimizantes que se le pueden poner a la maximización de la utilidad.

Pero con esto nos estamos anticipando. Teniendo en cuenta nuestra perspectiva de la racionalidad práctica, decidir entre las dos aspiraciones no parece constituir un problema. Parece evidente que las demandas de optimización de la utilidad deben prevalecer sobre cualquier demanda de optimización. La elección pone de manifiesto la preferencia. Si uno elige una estrategia y luego afirma que en lugar de su resultado esperado prefiere el resultado esperado de alguna otra estrategia que uno podría haber elegido, las preferencias de conducta y las preferencias expresadas por esa persona están en conflicto y sus utilidades (si se las identifica con los valores de esa persona) no quedan definidas. Esa persona no está en posición de obrar racionalmente. Si un individuo elige una estrategia y acepta que no debe preferir el resultado esperado de tal estrategia, el resultado de cualquier otra estrategia que pudiera haber elegido, la utilidad esperada de su resultado debe ser por lo menos igual a la utilidad del resultado de cualquier estrategia alternativa. De ahí que un individuo no puede elegir racionalmente una estrategia optimizadora antes que una estrategia maximizadora de su utilidad; al elegir una estrategia una persona muestra que, en la medida en que sus utilidades estén bien definidas, la considera maximizadora de su utilidad.

Si esta argumentación es correcta, la optimización no puede pues pasar por alto el hecho de que la maximización de la

utilidad es una exigencia de la elección estratégicamente racional, pero en el mejor de los casos puede servir como un fundamento para elegir entre varias estrategias maximizadoras de la utilidad. Y muchos teóricos de la elección racional consideran que la argumentación es sólida.[9] No podemos descartarla señalando que sólo la elección de una estrategia optimizadora sostiene las expectativas sobre las cuales se basa la elección. Utilizamos ese argumento para defender la elección de una particular estrategia maximizadora de la utilidad, pero un argumento similar no nos serviría para defender la elección de una estrategia que no sea maximizadora de la utilidad. Sólo puede haber una estrategia compatible con una expectativa particular, pero si, atendiendo a la expectativa, la estrategia no es maximizadora de la utilidad, sería pues plausible sostener que ni la elección de la estrategia ni la expectativa pueden considerarse racionales.

Aún no estamos en posición de replicar a ese rechazo de la racionalidad de la optimización. Más adelante trataremos de socavar la fuerza del enfoque que sostiene que la elección racional pone de manifiesto la preferencia, al mostrar que su alcance puede estar limitado por lo que en efecto es una metaelección, una elección sobre cómo hacer las elecciones. Pero, para concluir este capítulo, sencillamente reflexionemos sobre algunas de las consecuencias que trae apareado el hecho de exigir que un resultado estratégicamente racional esté en equilibrio, consecuencias que puede hacer más atractiva nuestra renuencia a aceptar directamente la identificación entre la racionalidad y la maximización de la utilidad.

3.2 El "dilema del prisionero", un cuento que se le atribuye a A. W. Tucker, constituye la ilustración clásica de las consecuencias adversas de la maximización de la utilidad.[10] Fred y Ed cometieron un grave delito (la fiscal del distrito no tiene dudas al respecto), pero parte de la prueba necesaria para asegurar una convicción es, desgraciadamente, inadmisible en el tribunal (la fiscal del distrito maldice a los reformadores de las leyes que dificultan aun más su tarea). Con todo ha logrado encerrar a Fred y a Ed y mantenerlos incomunicados entre sí. Ella los ha acusado de cargos que, aunque menores, siguen siendo graves y la fiscal confía en poder obtener su convicción por éstos. Los

visita entonces por separado y a ambos les hace el mismo cuento y la misma proposición: "Confiese que se ha descarriado y el crimen que cometió", le dice a cada uno, " y si su antiguo socio no confiesa, yo convenceré al jurado de que usted es un hombre reformado y su ex socio la encarnación del mal; el juez lo sentenciará a usted a un año de prisión y a él a diez. Si usted no confiesa y su antiguo socio lo hace, ya puede usted imaginar cuál será su destino. Y si ninguno de ustedes decide confesar, propondré que los condenen a ambos a dos años de prisión." "Pero, ¿que ocurre si ambos confesamos?", pregunta Fred (o Ed). "En ese caso", dice la fiscal del distrito, "dejaré que la justicia siga su curso natural. Se trata de un grave delito y estimo que los condenarán a cinco años a cada uno." Sin agregar nada más, la fiscal se retira y deja a Ed (o a Fred) en solitaria reflexión.

Lo único que le interesa a cada uno de los prisioneros es minimizar el tiempo que ha de pasar tras las rejas de modo que ambos representan la situación del modo siguiente:

		El	
Yo	confieso	confiesa 5 años a c/u	no confiesa 1 año a mí 10 años a él
	no confieso	10 años a mí 1 año a él	2 años a c/u

Y ambos razonan: "Si él confiesa a mí me conviene confesar, de lo contrario pasaré adentro diez años. Si él no confiesa y yo sí saldré en un año. Por lo tanto, haga lo que hiciese él, a mí me conviene confesar. En realidad (siendo ambos estudiantes de la racionalidad estratégica y de la teoría de los juegos) en esta situación hay un sólo resultado que está en equilibrio: el producto de la confesión mutua. No tengo dudas de la decisión que debo tomar".

Al día siguiente la fiscal de distrito (también ella estudiante de la racionalidad estratégica y de la teoría de los juegos) regresa a las celdas de Fred y de Ed con una sonrisa dibujada en la cara dispuesta a recoger ambas confesiones que sabe han de entregarle los prisioneros y que los mantendrá alejados de la mala senda durante cinco años; mientras que si Fred y Ed hubiesen mantenido las bocas cerradas (o se hubiesen negado a firmar la

confesión) hubieran podido regresar a sus ocupaciones favoritas a los dos años.

Este puede ser un resultado feliz de apuntar a una estrategia maximizadora de la utilidad mutua antes que a una estrategia mutuamente optimizadora. (En esta situación, cualquier resultado que se obtenga de la decisión —de un prisionero o de ambos— de no confesar es óptima; sin embargo, la simetría de la situación pone de manifiesto que una condición suficiente de la elección basada en la optimización exigiría que ninguno de los dos confiese.) Pero si nos abstraemos de este cuento en particular, reconocemos la seriedad del problema: en cualquier situación con una estructura paralela a la que afrontan Fred y Ed, los maximizadores de la utilidad supuestamente racionales se perjudican más de lo que podrían hacerlos los optimizadores supuestamente irracionales.

Otra ilustración nos mostrará hasta qué punto es desafortunado este problema.[11] Supongamos que hay dos naciones que por conveniencia (y aclaremos que cualquier aparente referencia a naciones reales es pura coincidencia) llamaremos *US* y *SU*. Estas dos naciones están comprometidas en una carrera armamentista, cuyos peligros ambas advierten, por lo tanto ninguna de las dos desea que se declare la guerra total. El desarme mutuo alejaría la amenaza de la guerra y no tendría (supongamos) ninguna desventaja (por ejemplo, una depresión de la economía); de modo que ambas naciones prefieren el desarme mutuo a la continuación de la carrera armamentista. Sin embargo, no hay modo de asegurar el acatamiento de un acuerdo de desarme y cada nación sabe que si es la única que se desarma, queda a merced de la otra, mientras que si es la única que conserva su armamento, puede convertirse en la potencia dominante del mundo. Cada una prefiere la carrera armamentista a pesar del peligro de guerra total, a quedar a merced de la otra y cada una prefiere la perspectiva de convertirse en la superpotencia al desarme mutuo. De ahí que, antes de firmar un acuerdo de desarme, cada nación representa las consecuencias que puede tener para sí cada conjunto de estrategias.

		SU	
		viola el acuerdo	lo respeta
US	viola el acuerdo	3° 3°	1° 4°
	lo respeta	4° 1°	2° 2°

La estructura de esta situación es precisamente paralela a la que afrontaban Fred y Ed. Por lo tanto el razonamiento también es paralelo. Haga lo que hiciere la otra nación, la violación del acuerdo maximiza la utilidad de la primera nación. El único resultado en equilibrio es la mutua violación. Ni hace falta decir que *US* y *SU* no han de firmar un acuerdo que no vale ni el papel en el que se lo escriba, sino que ambas limitan sus negociaciones o bien a acuerdos exigibles o a cuestiones en las cuales la aceptación mutua es (y ambas naciones reconocen que lo es) la alternativa más preferida de cada nación.

No cuestionamos la racionalidad de *US* y de *SU* en la situación en la que en realidad se encuentran. Para esas naciones la optimización mutua no es una alternativa viable a la maximización de la utilidad mutua. El mundo de las naciones es aquello que Hobbes llamó la condición natural de la humanidad y en esa condición natural no hay restricciones racionales al esfuerzo que hace cada persona o cada nación individual para maximizar su propia utilidad.[12] Pero por supuesto, esto es precisamente lo que está mal en el mundo de las naciones. La condición natural de la humanidad es una condición irracional porque condena a las personas o a las naciones a resultados no óptimos que, en situaciones del tipo de la del dilema del prisionero, no se diferencian mucho del desastre. Ninguna persona o nación individual es capaz de salirse unilateralmente de la condición natural de la humanidad. Pero todos podrían beneficiarse de terminar con un estado de cosas en el cual es racional interrelacionarse basándose únicamente en la maximización de la utilidad mutua. Lo que mostraremos es una exposición razonada y más estrictamente una exposición razonada maximizadora de la utilidad que nos permita reemplazar la "respuesta maximizadora de la utilidad" por una "respuesta optimizadora" al interpretar las condiciones *A*, *B* y *C* que rigen la racionalidad estratégica.

Notas

1 La teoría de los juegos se ha concentrado primariamente en este caso ideal. Véase la discusión en R. D. Luce y H. Raiffa, *Games and Decisions*, Nueva York, 1957, pág. 49.

2 Estas convenciones son las utilizadas normalmente en la exposición de la teoría de los juegos. Nuestro análisis no pretende ser aquí original.

3 Sobre la prueba véase J. F. Nash, "Non-cooperative Games", *Annals of Mathematics* 54, 1951, págs. 286-295.

4 Para una exposición de la posición esbozada aquí véase J. C. Harsanyi, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge, 1977, págs 114-115.

5 El enfoque de Harsanyi de la elección racional en tales situaciones abandona la racionalidad estratégica basada en las expectativas sobre las elecciones de los demás y se sitúa en una perspectiva "maximínima". Véase Harsanyi, *Rational Behavior*, especialmente las págs. 116 a 138. Alguien podría cuestionar la consistencia de la aceptación que hace Harsanyi de las estrategias maximínimas rechazando vehementemente el empleo que le da Rawls a los argumentos maximínimos en la teoría ética. Véase "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory", en *Essays on Ethics*, especialmente las págs. 39-40.

6 Véase F. Zeuthen, *Problems of Monopoly and Economic Warfare*, Londres, 1930, págs. 111-121.

7 Véase Harsanyi, *Rational Behavior*, págs. 133-135, 273-288; T. C. Schelling, *The Strategy of Conflict*, Cambridge, Mass., 1960, págs. 54-58, 83-118; y mi "Coordination", *Dialogue* 14, 1975, págs. 195-221.

8 El término usual es "Pareto-optimidad", usado por el economista italiano Vilfredo Pareto. Pero éste habla de la "ofelimitud máxima" de la que gozan los miembros de una colectividad. Véase su *Manual of Political Economy*, traducción de Ann S. Schwier, Ann S. Schwier y A. N. Page (comps.), Nueva York, 1971, pág. 261.

9 Como dice Harsanyi, "según nuestra teoría...las consideraciones de eficiencias sólo son válidas *dentro* de un conjunto [de estrategias] exigibles", págs. 127-128, *Rational Behavior*.

10 Véase Luce y Raiffa, pág. 94.

11 Esta ilustración ya resulta familiar en la bibliografía. Véase mi "Morality and Advantage", *Philosophical Review* 76, 1967, págs. 460-475.

12 Véase Hobbes, *Leviathan*, capítulo 13, pág. 63.

IV

El mercado: la libertad que procede de la moral

1. La economía es la disciplina pionera de la interacción racional. Y tal disciplina alaba un ideal de interacción libre de toda restricción. En dos breves pasajes de *La riqueza de las naciones*, Adam Smith formuló la declaración canónica de ese ideal:

> Al desechar...por completo todos los sistemas de restricción, el evidente y sencillo sistema de libertad natural establece su propio acuerdo. Cada hombre, mientras no viole las leyes de la justicia, queda perfectamente libre de perseguir su propio interés a su manera y de hacer competir su industria y su capital con lo de cualquier otro hombre...[1]

Y el resultado de esta competencia libre y natural armoniza la satisfacción del interés privado de cada persona con la satisfacción del interés social. Al generalizar partiendo de una instancia particular, Smith sostiene:

> Al preferir apoyar la industria nacional en lugar de la extranjera, el hombre sólo tiende a buscar su propia seguridad; y al dirigir tal industria de manera tal que su producto pueda ser el de mayor valor, tiende a lograr su propia ganancia y en este, como en muchos otros casos, ese individuo es conducido por una mano invisible a promover un fin que no formaba parte de su intención primaria. Y ni siquiera es esto siempre malo para la sociedad que no formara parte de su intención. Frecuentemente, al perseguir su propio interés, el individuo promueve el interés de la sociedad con más eficacia que cuando realmente intenta promoverlo.[2]

Smith caracteriza el "sistema de libertad natural" como un mercado perfectamente competitivo. La idea de semejante mercado esclarece nuestra comprensión de la interacción racional al revelar una estructura en la cual los intereses divergentes y aparentemente opuestos de los diferentes individuos se armonizan cabalmente. Concebido como un tipo ideal, el mercado perfecto, como ya lo hemos visto, garantiza la coincidencia del equilibrio y de la optimidad, y así su estructura constituye la verdadera antítesis del dilema del prisionero.

El mercado es un tema central de nuestra indagación. En realidad, en el capítulo 3 sostendremos que una de las cláusulas del acuerdo que permite establecer una moral racional sienta las bases para la elección de un mercado perfectamente competitivo, o una aproximación cercana, como una institución social, en aquellas circunstancias en las cuales un mercado semejante constituye una alternativa practicable. Pero obsérvese que lo que afirmamos es que el mercado debe elegirse allí donde corresponda, porque uno de los aspectos centrales de nuestra posición es que el mercado no siempre corresponde, que encarna una estructura muy especial de interacción que no puede abarcarlo todo. El mundo podría ser un lugar mejor si fuera, como lo han creído los economistas entusiastas del *laissez-faire*, un mercado perfectamente competitivo. Pero no lo es. Y, como mostraremos luego, precisamente porque el mundo no es un mercado la moral constituye una restricción necesaria para la interacción de las personas racionales.

La moral procede del fracaso del mercado. El primer paso para comprobar la veracidad de esta afirmación es mostrar que el mercado perfecto, si fuera posible, constituiría una zona moralmente libre, una zona en la cual no habría lugar para las restricciones morales. Al dejar que cada persona persiga libremente su propio interés a su manera el mercado satisface el ideal de *anarquía moral*. Y nosotros estamos más familiarizados con la anarquía como ideal político. Los anarquistas políticos nos ofrecen un panorama de una sociedad humana que no tiene ni necesita restricciones internas: una sociedad de personas pacíficas, productivas y solidarias cuyas interacciones están benditamente libres de toda autoridad o compulsión. El anarquista político intenta convencernos de que es posible un orden de asuntos humanos que no exija el artificio de la política. El anarquista moral nos ofrece un panorama comparable al de una sociedad humana que no tiene ni necesita restricciones internas: una sociedad formada por personas pacíficas, productivas y solidarias que sin embargo carecen de conciencia. El

anarquista moral intenta convencernos de que es posible un orden de los asuntos humanos que no exija el artificio más profundo de la moral.

En este capítulo determinaremos las pretensiones de la anarquía moral. Analizaremos el mercado perfectamente competitivo como una estructura de la interacción racional y mostraremos que su operación no debería necesitar restricciones a las elecciones que maximizan la utilidad individual. En realidad lo que mostraremos es por qué tales restricciones están en principio racional y moralmente injustificadas. Pero nuestra indagación revelará también los límites del mercado. De ese modo prepararemos el terreno para examinar luego una estructura completamente diferente de la interacción racional, que sí exige restricciones de la elección que maximiza la utilidad individual. Comprender la naturaleza, el papel que desempeña y los límites de una zona moralmente libre de los asuntos humanos nos ayudará a entender la naturaleza y la necesidad de una zona moralmente (y políticamente) restringida.

Al defender el "sistema simple de la libertad natural", el propio Smith observa que cada hombre es libre "mientras no viole las leyes de la justicia". Pero la justicia, como podrá recordar el lector, está conspicuamente ausente desde la explicación que hace Thomas Hobbes de la condición natural de la humanidad, explicación en la que "la Fuerza y el Fraude son las dos virtudes cardinales".[3] Y la ausencia de la fuerza y el fraude es esencial para que el mercado funcione, antes de que la mano invisible de Smith pueda realizar su beneficioso trabajo. Primero es necesario exorcizar la guerra que menciona Hobbes de cada hombre contra cada hombre. Y esto, como veremos luego, significa que el ideal de interacción libre que alababa Smith no es natural sino que es artificial y procede, para las personas racionales, de un marco de restricciones acordadas. Al entender el mercado perfecto como una zona moralmente libre tenemos que remontarnos a la moral que subyace en él y que lo antecede.

2.1 La idealización de la interacción representada por el mercado perfectamente competitivo incluye descartar tanto la incertidumbre circunstancial como el cálculo estratégico. La toma de decisiones del mercado, aunque compleja desde el punto de vista práctico, es del tipo más simple desde el punto de vista

lógico. Cada individuo elige paramétricamente, como si su acción fuera el único factor variable y considera las acciones de los demás como circunstancias fijas. Y cada uno elige como si conociera el resultado de sus posibles acciones. La médula de la teoría económica tradicional, que examina el funcionamiento del mercado, se preocupa pues por analizar la elección racional partiendo de un supuesto de certeza.[4]

No nos sumergiremos aquí en una completa descripción de los funcionamientos familiares del mercado, sino que solamente destacaremos los aspectos que son relevantes para nuestra argumentación. Los bienes del mercado son sus productos; los males son los servicios intermediarios utilizados en la producción. Cada persona desea consumir la mayor cantidad posible de productos, y la mayor cantidad posible de cada uno de ellos, aunque se supone que el deseo disminuye a medida que se aumenta el consumo. Cada persona quiere además brindar la menor cantidad de servicios intermediarios posibles y aquí se supone que el deseo de no brindar esos servicios disminuye al tiempo que también disminuyen la oferta de tales servicios. Además de ser el *locus* donde se dan la producción y el consumo, el mercado es el lugar del intercambio. Las condiciones necesarias para lograr la optimidad de la producción y el intercambio y para convertir los servicios intermediarios en bienes de consumo pueden encontrarse en cualquier estudio elemental de la teoría económica de modo que no vamos a detenernos aquí a examinarlas.

El mercado perfectamente competitivo supone que todos los productos y todos los factores de producción estén en manos privadas. De modo que el mercado queda especificado no solamente por las funciones de utilidad de aquellos que interactúan en él, que son las que determinan la demanda, y las funciones de producción que reflejan la tecnología existente, lo cual determina la oferta, sino también por una distribución inicial de los recursos que es la que le proporciona a cada persona su dote inicial. A los efectos de nuestro análisis, podemos definir a cada persona atendiendo a su función de utilidad y a sus recursos disponibles iniciales. Estos elementos fijan las preferencias y las facultades de la persona, las cuales son individualmente relevantes para la actividad que desarrolle en el mercado. La

distribución inicial de los recursos, y por lo tanto la definición de cada individuo que en parte depende de ella, ha de desempeñar un papel significativo en nuestro análisis y deberíamos observar desde el comienzo que en la perspectiva de la actividad del mercado, de la producción y del intercambio, es sencillamente una condición dada.

Decir que un recurso o un producto pertenecen a un individuo equivale a decir que esa persona puede utilizarlo a su gusto en el proceso de producción, intercambio y consumo. El derecho de posesión se extiende únicamente al uso de mercado, pero se extiende plenamente en ese uso. De modo que el mercado incluye la actividad enteramente libre de cada individuo, limitada únicamente por los recursos y los productos que le pertenecen, las funciones de producción que determinan las posibilidades de transformar los recursos en productos y las funciones de utilidad de los demás que determinan las posibilidades de intercambio. El supuesto de la propiedad privada puede dividirse en dos partes: los *recursos disponibles individuales*, y la *actividad del mercado individual libre.*

El mercado perfectamente competitivo supone una segunda forma de privacidad en el consumo de todos los productos. Cada bien unitario forma parte de la función de utilidad de una sola persona. Esto tiene una implicancia doble. Por un lado, todos los bienes deben ser estrictamente privados en el sentido de que si una persona consume una unidad, ninguna otra persona puede consumir esa misma unidad. La comida es un bien puramente privado; lo que come una persona no puede ser comido por ninguna otra. Una autopista no es un bien puramente privado; muchas personas pueden utilizarla. Por otro lado, las funciones de utilidad deben ser estrictamente independientes; ninguna persona gana o pierde simplemente tomando utilidades de los demás. La utilidad de cada persona queda estrictamente determinada por los bienes que consume y los servicios intermediarios que suministra. En otras palabras, los conjuntos de bienes consumidos y de servicios suministrados son estrictamente distintos y la utilidad de cada persona queda determinada únicamente por la medida y la composición de su conjunto.

La restricción de los bienes privados tiene considerable

significación en la medida en que limita el alcance del mercado perfectamente competitivo. En efecto, la zona libre desde el punto de vista moral es la zona de los bienes privados. Pero, si bien este aspecto ha de tener su importancia en un análisis posterior, en el presente capítulo concentraremos la atención en la idea que exige que las funciones de utilidad sean independientes. John Rawls ofrece una feliz expresión para ilustrar esta idea: "se concibe a las personas como seres que no se interesan en los intereses de los demás".[5] En realidad la expresión es más restrictiva de lo necesario; el mercado sólo exige que se conciba a las personas como seres que no se interesan por los intereses de aquellos con los cuales realizan sus intercambios. Esta es la exigencia de "no-tuismo" de Wicksteed: mis preferencias no te incluyen, aunque pueden incluir a una tercera persona que no participe de nuestra interacción.[6] Por consiguiente, el consumo privado, como la propiedad privada, se divide en dos partes: los *bienes privados* y el *desinterés mutuo.*

La existencia de recursos o productos sin dueño o la presencia de bienes públicos o utilidades interdependientes puede dar lugar a las externalidades económicas. La ausencia de *externalidades* —aunque éstas puedan aparecer—es una condición adicional de la competencia perfecta. Una externalidad surge cada vez que un acto de producción o intercambio o consumo afecta la utilidad de alguna persona que no participa o que participa involuntariamente de esa interacción. Tal efecto puede ser, por supuesto, benéfico o dañoso; si es benéfico hablamos de una externalidad positiva o de eficiencia externa, si es dañoso, hablamos de una externalidad negativa o de ineficiencia externa.

Será útil dar un ejemplo de cada uno de ellos. Supongamos que un grupo de navieros se ponen de acuerdo para levantar un faro que ayudará a la navegación de sus naves. Gracias a la libertad del mar, cualquiera puede beneficiarse con la ayuda que proporcione la luz. Los que construyeron el faro no tienen ninguna forma de restringir el acceso al bien que proporcionan a aquellas personas que no compartieron los costos de la construcción. El efecto para aquellos que no participaron de la construcción del faro es pues una externalidad positiva.

Supongamos ahora que el dueño de una fábrica decide

descargar los residuos de su industria en la atmósfera, con lo cual aumenta la polución. Si cualquiera puede utilizar el aire a su gusto puesto que es un bien libre, no habría forma de exigirle al dueño de esa fábrica que compensara a los demás por los efectos malsanos que les produciría su manera de desahacerse de los residuos tóxicos. El efecto para cualquier persona que respire el aire contaminado es pues una externalidad negativa.

Obsérvese que en estos ejemplos la externalidad aparece a causa de la existencia de un bien libre: un bien que no pertenece ni a un individuo ni a un grupo de individuos sino que está disponible para el uso de todos. La existencia del mar como un bien libre le permite a cualquiera tomar ventaja de las mejoras que se hagan para la navegación. La existencia de la atmósfera como un bien libre le permite a cualquiera tomar ventaja de su capacidad de vaciadero. Al mostrar que esta libertad permite la aparición de las externalidades no estamos sugiriendo que el mar y el aire deban convertirse en posesiones privadas. Sin embargo estamos mostrando que impiden el funcionamiento eficiente del mercado.

Para que el mercado perfectamente competitivo funcione es esencial el emparejamiento marginal de la oferta y la demanda. Las externalidades perturban ese emparejamiento. El mercado tiende a disminuir la oferta de faros y a aumentar la oferta de polución. La explicación es directa. La demanda sólo es efectiva como disposición a pagar los costos de la oferta. Pero aquellos que se benefician de la disponibilidad de los faros tienen muy poco incentivo o ninguno para pagar sus costos. Puesto que los mares son libres es imposible canalizar la demanda y hacerle pagar un derecho de uso. Y así la demanda resulta ineficaz. Aquellos que ofrecen los faros lo hacen sólo en la medida de su propia demanda y no en la medida de la demanda adicional representada por los usuarios que no pagan. Por lo tanto, la oferta tiende a quedarse corta respecto de la plena demanda. Para utilizar la terminología del capítulo anterior, diremos que el resultado está en equilibrio —cada uno actúa de un modo maximizador de la utilidad atendiendo a las acciones de los demás—, pero no es un resultado óptimo —todos se beneficiarían si la demanda de faros pudiera canalizarse de modo tal que impulsara un aumento de la oferta—.

En el caso de la polución tenemos la situación opuesta. Aquellos que la producen, dado que el aire está disponible para que lo usen todos, deben pagar solamente sus costos directos y no los costos que la polución les provoca a los demás. De ahí que la falta de demanda no inhibe la oferta. Hay tanta oferta de polución que el costo marginal que tiene para quienes la producen equivale al beneficio marginal que les proporciona; los costos adicionales que les provoca a los demás no ejercen ningún efecto restrictivo sobre tal oferta. Una vez más el resultado está en equilibrio pero no es óptimo: todos podrían beneficiarse si se controlara la polución hasta el punto de que el costo marginal que les provoca a todos se equipare con el beneficio marginal que les proporciona a todos (aunque la posibilidad de que todos se beneficien depende del modo en que se asignen los costos y los beneficios).

Los recursos disponibles individuales y los bienes privados, la actividad del mercado libre, el desinterés mutuo y la ausencia de externalidades son los supuestos básicos de un mercado perfectamente competitivo. En él tanto la producción como el intercambio se desarrollan en situaciones de certeza. Las funciones de producción fijas, que representan una tecnología conocida, aseguran la certeza de la producción. Los precios conocidos del mercado, que varían según el conjunto de la oferta y la demanda y que son efectivamente independientes de la conducta de cualquier individuo, aseguran la certeza del intercambio. Si bien los precios están determinados por las preferencias de aquellos que participan de la interacción del mercado, sin embargo se les presentan a los actores como datos objetivos a los cuales deben ajustar sus actividades. Ya señalamos la aparición de esos valores comunes, universales y aparentemente objetivos, en la sección 4.2 del capítulo 2. El análisis marxista propone una palabra útil para esta apariencia de objetividad; dice que es una forma de *fetichismo.*[7]

La actividad libre en situación de certeza es suficiente para asegurar que el mercado se dirige constantemente hacia una condición de equilibrio. Y el economista ofrece ahora su demostración triunfal de que, dada una competencia perfecta, el equilibrio del mercado debe ser también óptimo: nadie puede medrar a menos que algún otro empeore su situación.[8] Hemos

visto que la aparición de las externalidades puede conducir a un equilibrio no óptimo, pero en una competencia perfecta, las externalidades están necesariamente ausentes. En el resultado del mercado perfectamente competitivo, nadie puede consumir más productos que los que le permiten los servicios intermediarios que presta, o no puede prestar menos servicios que los que corresponden a los productos que consume, a menos que otra persona consuma menos productos y brinde los mismos servicios o proporcione más servicios sin consumir más productos. De modo que el análisis del economista hace visible la mano invisible de Adam Smith. Al establecer las condiciones necesarias para que se dé el mercado perfectamente competitivo, mostramos aquellos rasgos que, si están presentes en la estructura de la interacción, aseguran que cada individuo, al buscar solamente su propia ganancia, promueva el interés de la sociedad pues ayuda a provocar un resultado óptimo mutuamente beneficioso, aun cuando esto no forme parte de su intención.

El hecho de que en el resultado de un mercado perfectamente competitivo coincidan el equilibrio y la optimidad es necesario para que podamos sostener que esa coincidencia proporcionaría una zona libre desde el punto de vista moral. En el equilibrio que resulta después de que se realizan todos los intercambios voluntarios, queda asegurada la ganancia individual; cada uno hace lo mejor que puede teniendo en cuenta las acciones de los demás. Si el resultado es óptimo queda asegurado el beneficio mutuo; cada uno hace lo mejor que puede teniendo en cuenta el rendimiento que reciben los demás. La ruptura del equilibrio, que resulta de la prosecución de la ganancia óptima individual es la fuente de las quejas que elevaron Hobbes y otros contra la condición natural de la humanidad. El mercado perfecto supera esa falla y al hacerlo da lugar al panegírico de Smith a la mano invisible. Pero no debemos sencillamente dejar de lado los otros rasgos del mercado que hemos señalado ya; pues no son meramente condiciones de la coincidencia entre el equilibrio y la optimidad. Aunque esa coincidencia ponga de manifiesto la plena racionalidad estratégica de la interacción del mercado, para sostener que el mercado ideal sería una zona libre desde el punto de vista moral, el supuesto de la actividad libre es independientemente necesario. Por lo tanto debemos preguntarnos qué es la actividad libre.

2.2 Consideremos la actividad de un ser solitario, de un Robinson Crusoe, como lo llamaremos. Crusoe es libre de utilizar sus habilidades del modo que mejor satisfaga sus preferencias, dadas las circunstancias externas en las que se encuentra. Sus habilidades, por supuesto, limitan la extensión de la satisfacción de sus preferencias. Sus preferencias afectan el empleo que nuestro personaje hace de sus habilidades —no limitando lo que Crusoe puede hacer, sino lo que él considera que vale la pena hacer—. Pero ninguna de esas limitaciones restringe la libertad de Crusoe. Una persona es libre en la medida en que es capaz —sin interferencias—de dirigir sus habilidades al servicio de sus preferencias. Puesto que Crusoe puede hacer esto, disfruta del beneficio pleno de sus esfuerzos; consume lo que produce. Y puesto que puede detener sus esfuerzos productivos cuando el costo de una producción adicional excede el beneficio de consumir el bien producido, sólo puede culparse a sí mismo si los costos y los beneficios adicionales de su actividad no son iguales. En la medida en que Crusoe sea una persona racional, el valor que tiene para él el último bien unitario que produce debe ser igual al costo que le provoca el último servicio intermediario unitario que ofrece por su producción.

El efecto del mercado perfectamente competitivo es asegurar, en el nivel interpersonal, la misma libertad de la que goza Robinson Crusoe en su soledad y al mismo tiempo hacer posible la cantidad enormemente ampliada de los beneficios que brindan el intercambio y la división del trabajo (y también, a largo plazo, la inversión). Como ya lo hicimos notar, el mercado perfecto reduce la interacción estratégica a la elección paramétrica; las funciones de la producción y la demanda, atendiendo a las cuales cada individuo elige el curso de acción que maximice su utilidad, son circunstancias fijas. De modo que, si bien el uso que cualquier individuo particular pueda hacer de su talento y la satisfacción que ese individuo pueda obtener de sus preferencias dependen de las consideraciones de la oferta y la demanda sumadas, y por lo tanto de los talentos y las preferencias de todos aquellos que participan de la interacción del mercado, nadie está en posición de controlar los términos de la interacción correspondientes a los demás.

Cada persona es pues un Robinson Crusoe, hasta en el

mercado. Por supuesto, la existencia de otras personas y las posibilidades de intercambio afectan los límites de la conducta de cada persona. No solamente sus propias habilidades, sino también las de los demás, limitan la satisfacción que cualquier individuo puede darles a sus preferencias. Y la afectan de dos maneras diferentes. Primero, la capacidad productiva conjunta de una sociedad está limitada por las habilidades de sus miembros. Y segundo, la retribución que cada persona puede esperar por el uso de sus habilidades en la provisión de los servicios intermediarios depende de las relaciones de sustitución entre sus habilidades y las habilidades de los demás. Si la persona promedio se eleva a las alturas de Aristóteles, mi talento filosófico no sería un producto fácilmente negociable.

Además no sólo las propias preferencias, sino también las preferencias de los demás, afectan el empleo que uno hace de sus habilidades, pues limitan lo que cada individuo considera valioso hacer. El talento que tiene Wayne Gretzky para desplazarse sobre el hielo le rendiría un escaso beneficio si no existiera un público al que le interese el hockey. Pero esas limitaciones constituyen la analogía interpersonal de las limitaciones que experimenta Robinson Crusoe en la soledad de su isla. Si su libertad no está restringida por sus limitadas habilidades ni por su interés particular, tal libertad ha de permanecer irrestricta cuando Crusoe abandone la isla y comience a relacionar sus habilidades y sus intereses con los de los demás.

Robinson Crusoe goza del beneficio pleno de su trabajo. De manera similar, aquellos que participan de la interacción en el mercado gozan del beneficio pleno de sus tareas. En condiciones normales, el valor del rendimiento total del mercado es igual a la suma de los valores marginales con los que cada persona contribuye a tal rendimiento: la suma de las diferencias marginales que hace cada persona. En los intercambios libres del mercado, cada individuo puede esperar una retribución de un valor equivalente a su contribución.[9] Por lo tanto, el beneficio que recibe cada uno, o el valor de los bienes que cada uno es capaz de consumir, es igual a la contribución que hace, o a la diferencia marginal que agrega al valor del producto total. Y como Robinson Crusoe en la soledad de su isla, en el mercado cada persona sólo puede culparse a sí misma si los costos y los

beneficios que le provoca su actividad no son equivalentes. En la medida en que ese individuo es una persona racional, el valor que tiene para él el último bien unitario que produce u obtiene en el intercambio es igual al costo que le provocó el último servicio intermediario unitario con el cual contribuyó a la producción general o el último bien unitario que ofreció en el intercambio. De modo que el mercado proporciona un resultado óptimo en el cual las retribuciones que perciben los individuos son iguales a las contribuciones que hicieron tales individuos. Y esto ocurre no gracias a una intención deliberada de los individuos que participan de la interacción del mercado, cada uno de los cuales sólo busca maximizar su propia utilidad, sino como resultado de la estructura de interacción que brindan las condiciones de una competencia perfecta.

3.1 Los economistas partidarios del *laissez-faire* consideran que la coincidencia del equilibrio y la optimidad en el resultado de la actividad desarrollada en el mercado libre justifica la libertad de cada persona de perseguir su propio interés. Cualquier restricción a esa libertad, cualquier interferencia en el natural funcionamiento del mercado, no les parece ni racional ni moralmente aceptable. Por supuesto, existen muchas formas de interferencia que pueden impedir que el resultado del mercado sea óptimo, es decir que dejan a ciertas personas peor de lo que deberían quedar, debido a las retribuciones que reciben otros. Pero aun cuando pudiera conservarse la optimidad, la interferencia puede incluir una redistribución de los productos que deje a algunas personas en mejor situación y a otras en peor situación de lo que hubieran quedado si la interacción hubiese sido plenamente libre. Los defensores del mercado preguntan: ¿qué podría darles a los beneficiarios derechos sobre esas ganancias obtenidas a expensas de sus semejantes? Y ¿qué podría resarcir a estos últimos de sus pérdidas? La única actividad que afecte el mercado que podría ser o bien racional o bien justa es simplemente preservarlo, protegerlo de la fuerza o el fraude, esas dos virtudes cardinales de la condición natural de la guerra que mencionaba Hobbes y que son los grandes enemigos de la interacción libre y mutuamente beneficiosa.

El argumento que podrían presentar los economistas par-

tidarios del *laissez-faire* para terminar con las restricciones políticas es también claramente un argumento para terminar con las restricciones morales. Si cada individuo no es moralmente libre de perseguir su propia ganancia, algunos deben beneficiarse a expensas de otros. Mientras los primeros pensadores vieron en la búsqueda desenfrenada de la utilidad individual la principal fuente de conflicto de las relaciones humanas, los defensores del mercado ven en ella la base de la verdadera armonía que resulta de la plena satisfacción simultánea de las preferencias que mide la utilidad. Se le dice al moralista tradicional que sus servicios son innecesarios.

Puede parecer evidente que la interacción del mercado es racional. El hecho de que en el resultado coincidan el equilibrio y la optimidad asegura que las condiciones de una elección racional enumeradas en el capítulo anterior queden satisfechas, siempre que supongamos que una respuesta racional a las elecciones de los demás es maximizadora de la utilidad u optimizadora. De modo que el argumento de los defensores del *laissez-faire* parece necesitar que se afirme que si la elección es tanto maximizadora de la utilidad como optimizadora, debe ser también moralmente válida. Pero una interpretación más profunda de ese argumento, que sin embargo apoyamos, es que éste se basa en la pretensión de que la moral no puede aplicarse a la interacción de un mercado que cumpla con las condiciones de la competencia perfecta. La elección no es correcta o incorrecta desde el punto de vista moral, porque la coincidencia de la maximización de la utilidad y de la optimidad en la interacción libre hace que las restricciones que establece la moral no sean ni necesarias ni racionales, restricciones que nos permiten distinguir entre las elecciones correctas y las elecciones incorrectas. Las restricciones morales sólo aparecen en la brecha abierta por el conflicto creado entre las dos propiedades de la racionalidad, cuando el beneficio mutuo no queda asegurado por la prosecución de la ganancia individual. Determinamos que un resultado es correcto o incorrecto solamente cuando maximizar la utilidad de un individuo, atendiendo a las acciones de los demás, impide maximizarla, atendiendo a las utilidades de los demás.

Defenderemos esta perspectiva partiendo de una concep-

ción del ser humano que lo considera como un actor individual racional (o potencialmente racional) y una concepción de la moral que la considera una restricción a la búsqueda individual de la utilidad. Pero queremos aclarar que nuestra argumentación no implica una defensa del *laissez-faire* en la práctica económica. Pues sus tradicionales defensores suponían —y nosotros no— que el mercado perfectamente competitivo se da, o casi se da, o por lo menos podría darse, en la mayor parte de nuestras transacciones económicas. Estos pensadores no advirtieron la inevitable presencia de las externalidades en casi todos los contextos, y si la advirtieron, subestimaron su magnitud. De ahí que sacaron conclusiones políticas de su posición normativa básica cuando en realidad éstas exigen premisas empíricas que parecen falsas, o en el mejor de los casos, dudosas. Lo que nos interesa es mostrar que en la interacción ideal podría haber una zona moralmente libre y no defender su presencia en la mayor parte de nuestras actividades cotidianas.

Además, los defensores del *laissez-faire* no advirtieron la significación normativa de algunos de los supuestos del mercado. Y hemos señalado que el mercado supone un consumo privado y ésta es una cuestión de hecho. El hecho de que una persona consuma un bien unitario impide o no impide que otra persona lo consuma y las utilidades de una persona son o no son independientes de las utilidades de otra persona. Pero el mercado también supone la propiedad privada de los recursos y los bienes y ésta es una cuestión normativa. Se considera que los recursos disponibles iniciales de cada persona son algo dado, pero, ¿por qué debería considerarse así? Cada producto y cada recurso de la producción pertenecen a algún individuo que puede utilizarlos como más le plazca en la medida en que participen de la producción, el intercambio y el consumo, pero, ¿por qué deberían los individuos ejercer ese control?

Aparentemente, el supuesto de la propiedad privada presenta algunos interrogantes, no sólo referentes a la supuesta falta de moral de la interacción que se desarrolla dentro del mercado, sino referentes a la racionalidad y la moral que implica aceptar la estructura del mercado para desarrollar la interacción. Y esos interrogantes son particularmente apremiantes si se tiene en cuenta que en cualquier mercado hay muchos resulta-

dos óptimos posibles y que el único real está determinado por los recursos disponibles iniciales de los individuos que participan de la interacción, dadas sus preferencias y la tecnología de la producción. El resultado real le proporciona a cada individuo una retribución igual a su contribución, pero los servicios intermediarios que presta un individuo deben depender de los recursos que posea y por lo tanto de sus dotes. Hemos supuesto que se puede definir a cada persona por su función de utilidad y por sus recursos disponibles. Pero la segunda parte de esta definición puede parecer completamente arbitraria, y cualquier otra definición diferente de cada persona —si el resto de los datos sigue siendo el mismo— conduciría a una diferente distribución final de los bienes.

Al responder a estas preguntas, distinguimos entre las *condiciones* del mercado y su *operación*. La operación del mercado consiste en convertir una situación inicial especificada atendiendo a los recursos disponibles de los individuos en un resultado final especificado atendiendo a una distribución de los bienes o los productos entre esos mismos individuos. Puesto que el resultado del mercado está en equilibrio y es óptimo, se demuestra que su operación es racional y puesto que ese producto procede de la actividad libre de los individuos, afirmamos que su racionalidad no da lugar a la evaluación moral. Dada la situación inicial del mercado, su resultado sólo puede estar plenamente justificado.

Pero ni la operación del mercado ni sus resultados pueden mostrar, ni siquiera tratar de mostrar, que su situación inicial también es aceptable ya sea desde el punto de vista racional, ya sea desde el punto de vista moral. De modo que en el argumento que vamos a presentar sobre la inaplicabilidad de una evaluación moral a la operación del mercado no hay nada que pueda mostrar que los recursos disponibles iniciales mismos no plantean cuestiones morales. En su análisis de la optimidad de la competencia perfecta, David Winch señala lo siguiente: "El sistema perfectamente competitivo depende de la posesión privada de los recursos de la producción", de modo que "la esencia del sistema es que la justicia distributiva es un atributo de lo que se aporta al sistema y no de lo que se obtiene de él".[10] Estamos completamente de acuerdo con esta afirmación. Cual-

quier defensa de los arreglos particulares del mercado debe incluir una defensa de los aportes de los recursos disponibles particulares. Pero, como afirma Winch, "la aceptación de una distribución inicial de los recursos nos impide hacer juicios de valor sobre la distribución de las utilidades". La operación de un mercado no puede en sí misma dar lugar a ninguna cuestión evaluativa. Los resultados del mercado son justos si, y por supuesto sólo si, proceden de condiciones iniciales justas. En los próximos capítulos veremos qué es la norma de lo justo.

3.2 Ahora dirigiremos nuestros esfuerzos a defender la idea de que el mercado perfecto crearía una zona moralmente libre para la actividad individual, partiendo de nuestra concepción inicial de la moral que la define como una restricción imparcial a la búsqueda directa de la utilidad individual. Nos preguntamos si la operación del mercado exhibe cualquier forma de parcialidad que, desde un punto de vista moral, podría justificar la restricción necesaria para superarla. Puesto que hemos defendido ya la posición de que la interacción del mercado es racional, si comprobáramos que es parcial, deberíamos haber establecido, por lo menos *prima facie* el conflicto que existe entre la moral y las exigencias racionales. Negaremos que exista tal conflicto. Pero no por ello debemos sencillamente suponer que, porque la interacción del mercado es racional y la moral no se opone a la racionalidad, el mercado es por lo tanto moralmente libre. Si nuestra conclusión no está obligada a hacer una petición de principio, debemos argumentar independientemente de nuestra concepción que considera que la moral es una restricción imparcial.

Quizá la mejor manera de decidir si la interacción del mercado exhibe alguna parcialidad es preguntarse si cualquier individuo puede afirmar razonablemente que la operación del mercado lo afecta de algún modo diferencial, si lo hace favorable o desfavorablemente, o si esa persona puede presentar cualquier queja u objeción razonable contra el resultado del mercado. Damos por sentado que ningún individuo puede objetar razonablemente ningún resultado al que él mismo haya contribuido como un Robinson Crusoe. Como en este momento lo que nos interesa es el funcionamiento del mercado, también damos por sentado, por lo menos inicialmente, la definición del indivi-

duo que da el mercado. Después, en esta misma sección, determinaremos qué parte de esa definición exige una defensa adicional, para que aparezcan como moralmente aceptables no sólo el funcionamiento sino también las condiciones del mercado.

Por lo tanto, primero señalamos que el supuesto de *actividad libre* asegura que ninguna persona esté sujeta a ninguna forma de compulsión, ni a ninguna forma de limitación que no afectara ya sus acciones como individuo solitario. Cada uno elige por sí mismo lo que produce, lo que intercambia y lo que consume, ante una tecnología de producción conocida y un conjunto de precios conocidos, que ni ese individuo ni ningún otro puede afectar. Si el mercado estuviera destinado a producir un resultado, aunque óptimo y en equilibrio, y permitiera que algún individuo o grupo de individuos ejerza una compulsión sobre otros, haríamos bien en suponer que el mercado afecta a las personas de manera diferente y tendríamos que insistir en señalar la necesidad de las restricciones morales que permitan superar esa parcialidad. Pero, teniendo en cuenta la libertad de la interacción del mercado, introducir restricciones morales sería introducir la compulsión misma que suponemos que la moral intenta contrarrestar. Puesto que el mercado no produce nada que deba ser corregido, introducir un artificio correctivo sería en sí mismo una fuente de parcialidad.

En segundo lugar, la *ausencia de externalidades* asegura que ningún individuo resulte afectado, ni beneficiado ni perjudicado, por ninguna actividad del mercado de la cual ese individuo no eligió formar parte. Cada uno no sólo actúa libremente, sino que también interactúa libremente. De modo que ninguno de los miembros que participan de la interacción está en la posición de jugador líbero o de parásito. El jugador líbero obtiene un beneficio sin pagar su costo o parte de su costo. El parásito, al obtener un beneficio, deriva el costo o parte del costo hacia otra persona. Recordemos nuestros ejemplos anteriores. Los navieros cuyas embarcaciones obtienen una ventaja para navegar de un faro, aun sin haber contribuido ni a su construcción ni a su mantenimiento, son jugadores líberos. Aunque no derivan los costos hacia los demás, obtienen una ganancia sin pagar ninguno de los costos que exige obtener esa

ganancia. El dueño de la fábrica que lanza los residuos gaseosos y aumenta la polución de la atmósfera sin compensar a aquellos que sufren las consecuencias es un parásito que deriva parte de los costos de sus actividades a los demás. Y obsérvese que continuaría siendo un parásito aun cuando los demás le pagaran para que no aumente la polución, pues entonces el dueño de la fábrica estaría cargándoles a los demás los costos de un método alternativo de desembarazarse de sus residuos tóxicos.

El jugador líbero no empeora con su conducta la situación de los demás. Aquellos que erigen y mantienen el faro maximizan sus utilidades al hacerlo. Pero el jugador líbero obtiene un beneficio del que no gozan los demás, de modo que el resultado lo afecta de un modo diferente favorablemente. El parásito empeora directamente la situación de los demás. Su actividad puede parecer más evidentemente injusta que la del jugador líbero. En realidad, viola directamente las condiciones de la actividad libre al forzar a los demás a participar con él de la interacción de un modo que ellos no eligieron. Y se beneficia de manera diferente. Pero aunque el parásito está doblemente en falta, ambos socavan la imparcialidad que debe existir en la interacción, por la cual cada uno paga los costos de los beneficios que obtiene.

Igualar la renta con la contribución marginal asegura precisamente esa imparcialidad. Pues no solamente hace que cada individuo participe libremente de la interacción con los demás, sino que asegura que cada uno se beneficie de y únicamente de la contribución que hace. A falta de una alternativa mejor, un individuo puede elegir libremente participar de una interacción diferentemente desfavorable con los demás. Viernes podría estar en mejor situación como esclavo que como un Robinson Crusoe; y si tuviera que elegir entre la esclavitud y la soledad, podría optar por la esclavitud. Pero Viernes no se beneficia de la interacción en una medida equivalente a su contribución. Las reglas del mercado descartan todas esas formas atípicas de interacción. A pesar de la propaganda que sostiene lo contrario, la esclavitud del salario no es ni puede ser un rasgo del mercado perfectamente competitivo. Este es un tema al cual regresaremos en la última sección de este capítulo en donde examinaremos una objeción marxista a nuestra defensa del mercado.

En tercer lugar, la *optimidad* del resultado del mercado asegura que ninguna alternativa (que no resulte peor para todos) beneficie a algunos a expensas de otros. Pues cualquier alternativa debe implicar la disminución de la utilidad de alguna persona y podemos considerar que esto es o bien una pérdida de los beneficios o bien un aumento de los costos. En el primer caso, la persona no ha de recibir algún beneficio cuyo costo pagó en la interacción libre; en el segundo caso, la persona debe pagar algunos costos adicionales además de los que serían suficientes en la interacción libre para cubrir todos los beneficios. En ambos casos, el interés de la persona se perjudica. Y esa persona puede quejarse razonablemente si alguna otra goza de un aumento de los beneficios o de una reducción de los costos, de que esa otra persona se benefició a sus expensas.

Pero es necesario calificar este último punto. El mercado le proporciona a cada persona una retribución por los servicios intermediarios que esa persona ofrece. Pero esa retribución puede incluir dos componentes diferentes. Primero, y necesariamente, el precio pagado en el mercado por un servicio intermediario debe cubrir el costo de su oferta. Cada persona que pone su capacidad y sus materiales al servicio de la producción y el intercambio, recibe una retribución que cubre los costos de desarrollar y mantener esa capacidad y obtener esos materiales. Si no fuera así, la interacción no sería voluntaria. Si cada individuo no recibiera el costo de su oferta debería pues ofrecer únicamente sus servicios intermediarios ante la amenaza de penas coercitivas. Pero en segundo lugar, en la medida en que ciertos recursos formen parte de una oferta fija —en la medida en que ciertas habilidades o ciertos materiales no puedan sustituirse libremente en el mercado— existe la posibilidad de que la demanda de los bienes producidos gracias a esas habilidades y con esos materiales den lugar a un precio de mercado para su utilización que incluya el rédito.[11] El que recibe el rédito se beneficia de la escasez de los factores que controla —una escasez que, desde su punto de vista, es enteramente accidental, puesto que depende, no de la naturaleza intrínseca de los recursos, sino de la relación que existe entre ellos y los recursos controlados por los demás—. Ese individuo recibe más de lo que necesita para aportar sus recursos al mercado; el rédito

es por definición, una retribución superior al costo de lo que se ofrece.

Si el resultado del mercado incluye el rédito, su optimidad no asegura directamente que ninguna alternativa (que no sea peor para todos) beneficie a algunos individuos a expensas de otros. Mientras cada individuo reciba el costo total que tienen para él los servicios intermediarios que ofrece y mientras cada individuo pague también el costo total de los servicios intermediarios que obtiene, aparentemente nadie se beneficia a expensas de otro. Si el superávit representado por el rédito va hacia quien controla un recurso escaso, o se distribuye de algún otro modo es algo que parecería haber quedado racional y moralmente abierto por todo lo que dijimos hasta ahora. Volveremos a examinar la cuestión del rédito al considerar el establecimiento del mercado perfectamente competitivo como uno de los términos del acuerdo realizado entre las personas racionales. Por ahora, debemos advertir que nuestro argumento destinado a apoyar la afirmación de que el mercado perfecto crea una zona moralmente libre, no parece aplicable a aquellas transacciones del mercado que incluyen el rédito. Con todo, podemos llegar a la conclusión de que la actividad libre, la ausencia de externalidades y la optimidad juntas son suficientes para derrotar todos los cargos de parcialidad que puedan hacérsele a la operación del mercado y aseguran que, dejando de lado el rédito, cualquier alternativa al resultado del mercado o bien dejaría a todos en peor situación, o bien sería parcial, pues permitiría que algunos se beneficien a expensas de otros. La idea de una zona moralmente libre tiene sustento.

O por lo menos, está sustentada para aquellas personas definidas por sus funciones de utilidad y sus recursos disponibles. Digamos que esas dos características constituyen la *identidad de mercado* de una persona; por lo tanto en la medida en que un individuo se identifique con su identidad de mercado ha de considerar, por supuesto, que el mercado es imparcial. Ahora bien, un individuo puede identificarse fácilmente con su función de utilidad, que es la medida de sus preferencias consideradas, pero puede no identificarse con sus recursos disponibles, pues puede considerarlos puramente arbitrarios. En realidad, un individuo puede rechazar la idea misma de dotes y afirmar que los supuestos del mercado no tienen justificación.

No es sorprendente que en las sociedades de mercado el proceso de socialización aliente a cada persona a identificarse con su *identidad de mercado*. Si llevamos esto al extremo, podemos imaginar una sociedad en la cual el resultado de la interacción (sea cual fuere) se considera óptimo y en la que cada uno de sus miembros está formado para encajar con ese resultado, es decir, para tener las preferencias y los recursos disponibles necesarios para que el resultado esté en equilibrio óptimo. Esta inversión de la sociedad de mercado, en la cual los individuos sirven al mercado antes que el mercado sirva a los individuos es una cuestión muy seria, tanto desde el punto de vista teórico como desde el punto de vista práctico. Volveremos a ella en los dos últimos capítulos de este libro y mostraremos que nuestra teoría moral exige rechazar tal idea.

No sostenemos que un individuo pueda identificarse razonablemente con su *market self*, sin embargo ese *self* (sí mismo) puede determinarse. Y queremos establecer que cada persona queda definida en parte por unos recursos disponibles apropiadamente establecidos. En realidad esta idea ya está supuesta cuando consideramos al individuo como un Robinson Crusoe y cuando hablamos de la libertad de que goza para poner sus habilidades al servicio de sus preferencias. Aunque las dotes naturales de Robinson Crusoe sean en cierto sentido arbitrarias, porque la persona no tendría ningún fundamento para quejarse si éstas fueran completamente distintas de lo que son, en otro sentido no son en absoluto arbitrarias. En la solitaria condición en la que se encuentra Crusoe, sus dotes son todo aquello de lo que él puede echar mano. De modo que esto le proporciona un determinado punto de partida para su conducta.

Extender esta concepción de la dote natural a situaciones que implican una interacción puede parecer problemático. Si suponemos que la dote de cada persona es aquello que puede utilizar, las dotes se superponen, de manera tal que como afirma Hobbes, cada persona " tiene derecho a todas las cosas, incluso al cuerpo de los otros".[12] Por otro lado, si suponemos que las dotes de cada persona son aquello que sólo ella utiliza, las dotes se desdibujan y nadie tiene derecho sobre nada, ni siquiera sobre su propio cuerpo. Pero, podemos resolver este problema modificando la concepción. Aun cuando Robinson Crusoe se encuentre

en realidad solo, podemos definir sus *dotes básicas* como aquello de lo que él puede hacer uso y que ningún otro puede usar en su ausencia. Y este concepto sí puede extenderse de la soledad al mercado. Las dotes básicas de cada persona son aquello de lo que la persona puede hacer uso simplemente en virtud de su presencia en una situación y que nadie puede usar en su ausencia. Esto comprende pues sus habilidades físicas y mentales. Y así se relaciona con la libertad de que gozan tanto Robinson Crusoe como aquellos que participan de la interacción del mercado.

Las dotes de una persona pueden incluir otros recursos de producción, y en realidad, para que sea posible la competencia perfecta, debe hacerlo. En el capítulo 7 mostraremos cómo pueden aparecer dotes semejantes y cómo hasta las dotes básicas pueden tener un fundamento racional. Pero para defender lo que nos interesa en este momento, tenemos que suponer únicamente que cada persona incluye en la concepción de sí misma tanto sus habilidades físicas como sus habilidades mentales. Y que las dotes básicas, aplicables tanto a Robinson Crusoe como a los hombres y mujeres que participan de la interacción del mercado, no son arbitrarias. Pues cada uno tiene una razón suficiente para considerar que la interacción de la que participa con sus semejantes es imparcial sólo mientras le brinde una retribución igual a los servicios que él aporta mediante el uso de sus habilidades.

3.3 Las actividades de un mercado perfecto muestran la despreocupación recíproca de los actores. Y esta concepción (de las personas que no se interesan por los intereses de los demás) es fundamental no sólo para la economía, sino también para una teoría moral. Porque coincidimos con la idea de Kant de que las restricciones morales deben aplicarse en ausencia de intereses dirigidos a los demás y a la idea de que en realidad deben aplicarse cada vez que los individuos puedan tener preferencias.[13] Pero diferimos de Kant en cuanto a reconocer un ideal de interacción humana libre de tales restricciones.

El supuesto de desinterés mutuo puede criticarse o bien porque se piensa que generalmente es falso, o bien porque sea verdadero o falso, se piensa que refleja una visión indebidamente desagradable de la naturaleza humana, destructora no sólo

de la moral sino también de los lazos que mantienen unida a cualquier sociedad humana. Pero una crítica semejante podría malinterpretar el papel que le cabe a ese supuesto. Por supuesto, las personas muestran interés por las demás, pero ese interés suele ser propiamente particular y parcial. No es irrealista ni pesimista suponer que más allá de los lazos de sangre y amistad, que son necesariamente de un alcance limitado, los seres humanos muestran pocos sentimientos positivos de solidaridad. Cuando terminan las relaciones personales, sólo queda un débil interés negativo, que se manifiesta quizás en una predisposición general a abstenerse de utilizar la fuerza y el fraude si los demás hacen lo propio y en una predisposición particular a ofrecer ayuda en situaciones extremas, por ejemplo, en el deseo de ayudar a las víctimas de los desastres, aun a un costo mayor que el que hubiera sido necesario para evitar el desastre. Pero este interés limitado es plenamente compatible con la idea de que cada persona se preocupa por sí misma en los asuntos corrientes de la vida y extiende una mano solidaria a amigos y parientes que a su vez también se la extienden a él.

Al analizar la posibilidad y las características de las relaciones morales y sociales en el caso limitado pero importante desde el punto de vista práctico en el cual los lazos particulares entre las personas o bien están ausentes o bien, si están presentes son demasiado débiles para sustentar tales relaciones, consideramos que las funciones de utilidad son independientes. Uno de los problemas que afronta la mayor parte de las sociedades humanas es la ausencia de cualquier forma de interacción efectiva y mutuamente beneficiosa entre personas que no estén vinculadas por algún lazo particular. De ahí que sean tan importantes los sistemas de parentesco; un individuo coopera con un pariente porque lo une a él un lazo de confianza, de certeza en cuanto a que aquellos con los cuales participa de una interacción no han de desconocer sus intereses. Pero las demás personas son enemigos potenciales; la ausencia de cualquier lazo determinado implica una falta de interés que se considera como el obstáculo que imposibilita cualquier relación cooperativa. Por ello, el tráfico de mujeres (para utilizar el rótulo que le pone Gayle Rubin a la práctica documentada por Lévi-Strauss) ejemplificada particularmente en los matrimonios

exogámicos de las sociedades patriarcales, es necesario para crear una red de rehenes que impida la erupción de la hostilidad latente entre grupos que no son parientes.[14] La distinción fundamental entre "nosotros" y "ellos", entre los hermanos de sangre y los extraños, ha limitado el alcance de la cooperación (y ha contribuido al sojuzgamiento de la mujer) entre los miembros de gran parte de la humanidad. Invocamos el supuesto del desinterés mutuo a fin de determinar si esa limitación es un mal ineludible de la condición humana.

La superioridad de la sociedad de mercado sobre sus predecesoras y rivales queda puesta de manifiesto en su capacidad de superar esa limitación y de dirigir el desinterés mutuo al beneficio mutuo. Si la interacción humana está estructurada por la condición de la competencia perfecta, no es necesario pues que exista ningún lazo entre aquellos que participan de tal interacción, salvo aquellos lazos que ellos mismos crean al perseguir sus propias ganancias. La impersonalidad de la sociedad de mercado, que ha sido objeto de una amplia crítica y a la que se ha acusado de la *anomia* y de la alienación de la vida moderna, es en cambio la base de la liberación fundamental que proporciona esa sociedad. Los hombres y las mujeres se han liberado de la necesidad de establecer más lazos particulares, sean éstos afectivos o coercitivos, a fin de mantener una interacción beneficiosa. La división entre parientes y extraños desaparece y no queda reemplazada por una nueva división según la cual los sujetos cooperan sólo mientras permanecen bajo la mirada vigilante del Gran Hermano. Contra el fondo del mutuo desinterés del mercado, pueden florecer relaciones humanas particulares de confianza y afecto sobre una base voluntaria. Aquellos que ansían las relaciones firmemente entrelazadas de otras formas anteriores de la sociedad humana, en realidad están tratando de huir de la libertad de elegir las personas por cuyos intereses han de manifestar algún interés.

Pero la liberación trae aparejados ciertos peligros. La sociedad de mercado libera a los individuos de los lazos afectivos particulares y de los lazos coercitivos. En las sociedades cuasi de mercado reales existe la tentación, puesta de manifiesto tanto en el pensamiento como en la práctica, de considerar que todo el mundo social cae dentro del alcance del mercado y que por lo

tanto está libre de toda forma de restricción. Por un lado, el ideal de una zona moralmente libre se extiende hasta abarcar toda la interacción social y por el otro, se interpreta que todas las relaciones humanas son un contrato basado en el interés individual que no se preocupa por el interés interpersonal. Por consiguiente, no se reconoce lo suficiente ni la necesidad de las restricciones morales y políticas ni la posibilidad de que existan vínculos humanos genuinamente afectivos.

Como ya lo hemos dicho anteriormente, una de nuestras ideas fundamentales es que la zona moralmente libre creada por el mercado sólo puede darse dentro de una estructura moral más profunda. Y mostraremos que la restricción moral no sólo es compatible con el desinterés mutuo, sino que en realidad es una exigencia racional, dados ese desinterés y las estructuras típicas de la interacción. Por lo tanto, proponemos defender la idea liberadora de una sociedad que no le impone ningún lazo afectivo a sus miembros y distinguiremos la libertad afectiva de la anarquía moral y política y una sociedad libre de un mercado que lo abarque todo.

La moral, como un sistema de restricciones racionalmente exigidas, es posible si tales restricciones se originan sencillamente en el entendimiento de que sirven para alcanzar de manera más efectiva los propios intereses, la mayor satisfacción de las propias preferencias, sean cuales fueren esos intereses y esas preferencias. De ese modo un individuo no puede huir de la moral aduciendo una falta de sentimiento o interés moral, o una falta de algún otro interés o actitud particular, porque la moral no posee tales bases afectivas. Hume creía que la fuente de la moral era la transmisión simpática de nuestros sentimientos de una persona a otra.[15] Pero Kant, justamente, afirmaba que la moral no puede depender de semejantes fenómenos psicológicos particulares por benévolo y humano que sea su efecto y por universales que puedan parecer tales fenómenos.[16] La afirmación kantiana de que la moral obliga independientemente de la naturaleza y el contenido de nuestros afectos, aparece, aunque sólo sea parcialmente, en la afirmación de que la moral se basa en el supuesto de desinterés mutuo.

Desde este punto de vista el mercado es necesariamente no moral. Porque en las operaciones de un sistema de competencia

perfecta no existen restricciones que las personas mutuamente desinteresadas deban o estén dispuestas a reconocer. Las restricciones serían no sólo insustanciales, sino también positivamente hostiles para la interacción de tales personas. La esencia del mercado es brindar un terreno de libertad racionalizado por su optimidad. Los primeros defensores del *laissez-faire* pueden ser perdonados por el entusiasmo con que recibieron esta idea, un entusiasmo que los llevó a aplicarla sin considerar hasta qué punto la competencia perfecta podía llevarse a cabo en los asuntos humanos. No suponemos que las sociedades de cuasi mercado verdaderas hayan hecho realidad, o hayan podido hacer realidad, el ideal de interacción libre y óptima. Y si bien podemos creer que esas sociedades que tendieron al ideal del mercado se acercaron más a su realización que aquellas otras que persiguieron objetivos distintos, en ausencia de una adecuada teoría alternativa ni siquiera podemos aseverar con confianza esta afirmación.

En los próximos capítulos analizaremos la interacción entre las personas en circunstancias que excluyen la armonía del equilibrio y la optimidad en la actividad libre que proporciona el mercado. Volveremos a los contextos estructurados por el dilema del prisionero y entonces, sin abandonar el supuesto del mutuo desinterés, encontraremos la forma de definir la moral como un conjunto de restricciones a la búsqueda de la utilidad máxima que toda persona racional debe reconocer. Pero el hecho de que mantengamos la idea del desinterés mutuo puede hacerle creer al lector que ignoramos el segundo peligro que trae aparejado la liberación de los lazos afectivos que provoca tal desinterés.

Ya nos hemos referido a ese peligro al señalar que en las sociedades de cuasi mercado reales hay una tendencia a considerar que en todas las relaciones humanas falta el interés interpersonal. Una sociedad que libera a sus individuos de la necesidad de desarrollar lazos afectivos a fin de comprometerse en la interacción social puede alentar a sus miembros a concebirse a sí mismos como átomos individuales no relacionados entre sí que carecen de la fuerza potencial para desarrollar vínculos genuinamente afectivos. El hombre del mercado es un maximizador de las utilidades basadas en el "no-tuismo"; puede

parecer que nuestra explicación le da al hombre del mercado una moral, pero le niega cualquier interés dirigido a los demás. Los lazos afectivos particulares que reconocimos al comienzo de esta subsección pueden constituir solamente un residuo de una imagen de los seres humanos que hemos rechazado.

La crítica no necesita negar que la individualidad, excepto como el privilegio de una elite numéricamente insignificante, ha aparecido sólo en las sociedades de cuasi mercado, para sostener que, sin embargo, esa individualidad se manifiesta autodestructiva, de modo tal que la liberación alcanzada por una sociedad que alienta el desinterés mutuo termina siendo alienación. Pero nosotros abrigamos la esperanza de llegar a un final más feliz y concluiremos nuestro estudio con un esbozo del individuo liberal que, sostendremos, goza de la afectividad libre que permite una sociedad basada en restricciones puramente racionales. Con esta observación promisoria abandonamos la cuestión del desinterés mutuo. Pero antes de dejar el tema de la interacción del mercado, examinaremos brevemente las fallas de dos teorías clásicas que rechazan la defensa del mercado como una zona moralmente libre que nosotros hemos ofrecido: el utilitarismo y el marxismo.

4.1 El utilitarista relaciona la moral con la sociedad como nosotros relacionamos la racionalidad con el individuo e identifica la moral como la maximización *colectiva*. El actor moral busca maximizar una medida que llamaremos *bienestar*, definida como una función estrictamente incrementadora de todas las utilidades individuales. En el presente análisis no necesitamos ninguna otra especificación precisa de la función del bienestar. El utilitarista generalmente la considera como una función aditiva de las utilidades individuales y supone que es una medida de la preferencia individual con significación interpersonal. Pero nuestra argumentación no se refiere en ninguna de sus partes a la posibilidad de semejante medida aditiva.

La maximización de la función de bienestar debe proporcionar un resultado óptimo. Pero no hay razón para esperar que ese resultado, al que denominaremos *bienestar óptimo*, coincida con el equilibrio óptimo que podría alcanzarse mediante la operación de un mercado perfectamente competitivo. En realidad, no

siempre ambos pueden coincidir, puesto que el bienestar óptimo depende sólo de las funciones de utilidad del individuo y de la tecnología de producción, mientras que el resultado del mercado depende también de la distribución inicial de los recursos de producción. Dotes iniciales diferentes producen resultados del mercado diferentes sin afectar por ello el bienestar óptimo.

Por consiguiente, el utilitarista se ve obligado a rechazar el resultado de la competencia perfecta. Si el principio utilitarista, que exige la maximización de la función de bienestar, se aplica a una sociedad de mercado, debe haber, según parece, una redistribución de los productos, diferente de la realizada para que exista un equilibrio de mercado. Pero como ya lo hicimos notar, una distribución semejante no puede satisfacer ni nuestro criterio de interacción racional ni nuestra concepción de moral. Al redistribuir los bienes, el utilitarista impone necesariamente costos a algunas personas además de a aquellas que pagaron completamente los beneficios que recibieron en la interacción libre y les confiere beneficios a otras que no pagaron sus costos. De modo tal que una sociedad que maximiza el bienestar crea jugadores líberos y parásitos.

La economía utilitarista y la economía del *laissez-faire* fueron productos de la misma revolución de la indagación social. Pero pronto esas dos teorías entraron en conflicto. El utilitarista, que comenzó siendo un aliado del economista al criticar las restricciones morales y políticas aún presentes en la sociedad de su época, se volvió contra la nueva sociedad de mercado y llegó a ser el abogado del estado providente: el estado que redistribuye los bienes para alcanzar el bienestar óptimo. El utilitarista no obtuvo gran éxito con esta defensa. Y no es sorprendente pues que la pretensión de compatibilizar el estado providente con el mercado libre sea en principio incoherente, pues incluye la aplicación simultánea de dos normas de evaluación incompatibles. El utilitarista busca el resultado, pero como afirma David Winch, "no podemos juzgar un sistema perfectamente competitivo por la equidad del resultado, porque la esencia del sistema consiste en que la justicia distributiva es un atributo de los aportes hechos al sistema y no de las retribuciones que se obtienen".[17]

Entonces, ¿por qué no habría el utilitarista de desviar su

interés de los rendimientos a los aportes? ¿Por qué no habría de intentar conciliar la maximización del bienestar con la aceptación de la operación del mercado aplicando el principio utilitarista para determinar la distribución inicial de los factores de producción? Esto es precisamente lo que queda implícito en la breve explicación que hace John Stuart Mill de la base utilitarista de los derechos. "Desde mi punto de vista, tener un derecho es pues tener algo cuya posesión la sociedad debe reconocerme y defender. Si un objetor me preguntara por qué la sociedad debe defenderla, yo no podría aducir ninguna otra razón que no fuera la utilidad general."[18] Podemos interpretar que Mill sostiene que los recursos disponibles de cada persona (aquellas cosas que esa persona puede utilizar mientras vea que encajan con la actividad del mercado y que, por lo tanto, la sociedad debe proteger de la intervención forzosa o fraudulenta de los demás) deben determinarse considerando qué distribución de los recursos ha de servir mejor a la utilidad general o, para utilizar nuestra terminología, sirve más para maximizar el bienestar.

En este breve pasaje Mill intenta establecer un profundo cambio conceptual de lo que entendemos por derechos (que para nuestra argumentación podemos identificar con los recursos disponibles). Reducir los derechos a la utilidad o el bienestar general puede parecer una forma de confundir la médula misma de la concepción del derecho: aquella facultad permitida a todos y a la que está condenado el bienestar. Los derechos parecen crear un espacio moral para cada individuo que la apelación a la maximización del bienestar le niega. Pero ahora podemos ver que ésta es una crítica demasiado simplista a la posición de Mill, pues no tiene en cuenta aquello que Mill trata de lograr.

Mill aplica la función de bienestar, o más específicamente el principio de máxima felicidad, directamente a los recursos disponibles de los individuos. Sostiene que las personas deberían gozar de aquellos derechos que las llevan a participar de la interacción de modo tal que el resultado del mercado es también el bienestar óptimo. Como ya observamos, al permanecer fijas la producción y las funciones de demanda, el equilibrio del mercado queda determinado por los recursos disponibles. Por lo tanto, debemos comenzar con la función de bienestar, aplicarla a determinar el resultado posible de la interacción del mercado

que maximizaría el bienestar y luego continuar trabajando para determinar el conjunto de derechos iniciales que podrían proporcionar ese resultado. Consideramos que la función de bienestar es la variable independiente y que las variables dependientes son los derechos individuales y los recursos disponibles.

El enfoque de Mill, si bien altera, no abandona lo que tradicionalmente se ha concebido como derechos. Los derechos de un individuo, sea como fuere que se los conciba, deben proporcionarle un espacio moral dentro del cual ese individuo pueda incorporarse a la actividad del mercado. El enfoque de Mill reconoce plenamente esto y lo admite. Al resolver el problema de relacionar las dotes iniciales con el bienestar óptimo, no podemos continuar entremetiéndonos con esas dotes o con la interacción del mercado que procede de ellas aplicando una vez más el principio utilitarista al resultado del mercado. Respetamos la afirmación de Winch referente a que el sistema perfectamente competitivo no debe juzgarse por sus rendimientos. Si pretendemos afirmar que el resultado satisface las exigencias tanto de la interacción racional como de la maximización del bienestar debemos admitir que la interacción procede de los derechos asignados.

Ahora el estado providente aparece no como el agente de la redistribución que perturba el resultado del mercado racional en nombre de un ideal moral irracional, sino, antes bien, como el factor determinante de uno de los hechos dados en la actividad del mercado. No hay una base para criticar la actividad del estado, si éste se limita a determinar los recursos disponibles, en el marco del mercado mismo o desde el punto de vista de los individuos que participan de la interacción del mercado. El estado no interfiere con la operación del mercado sino que sencillamente establece sus condiciones. Observamos al concluir la sección 3.1, que los arreglos del mercado son justos si y sólo si son el resultado de condiciones iniciales justas. El utilitarismo establece las normas de esa justicia o imparcialidad. La maximización del bienestar queda así conciliada con la operación del mercado perfectamente competitivo.

Pero antes de ponernos a alabar el utilitarismo en esta feliz perspectiva, debemos prestar atención a ciertas complicaciones que se presentan al tratar de determinar los recursos disponi-

bles que no tienen ninguna equivalencia en el sencillo enfoque de los derechos de Mill, pero que se hacen evidentes cuando pensamos en los demás elementos que contribuyen al resultado del mercado: las funciones de producción, que reflejan la tecnología, y la función de demanda (o las funciones de utilidad individual en las cuales se basa). En una economía cambiante y dinámica hay un cambio continuo y esencialmente imprevisible de las funciones de producción. La tecnología no es estática y la competencia del mercado acelera su crecimiento. Esto implica que una asignación fija de los recursos disponibles no ha de continuar con el tiempo proporcionando el mismo resultado del mercado. Los cambios tecnológicos alteran la productividad de los recursos de modos imprevisibles y de ahí que alteren los derechos finales a los beneficios. Una distribución de los recursos que, contando con una tecnología dada, proporciona beneficios relativamente equitativos a las personas que tienen diferentes recursos disponibles, al presentarse cambios tecnológicos, ha de proporcionarles a dichas personas beneficios altamente inequitativos. Y no es probable que ambos resultados puedan maximizar el bienestar.

Podemos decir que ciertos tipos de cambios tienen efectos característicos, aun cuando otros tipos sean menos predecibles. Un ejemplo, analizado por Winch, puede aclarar la visión del lector no familiarizado con la bibliografía económica.[19] El aumento de la productividad de la agricultura disminuye los niveles de utilidad de los agricultores, de los ganaderos y de los trabajadores agrarios a causa de la relativa falta de elasticidad de la demanda de los productos agrícolas. Por otro lado, el aumento de la productividad industrial tiende a aumentar los niveles de utilidad de los propietarios y obreros de las fábricas a causa de la elasticidad relativamente alta de la demanda de los productos industriales. En la medida en que los individuos lleguen a relacionar los niveles de utilidad o la distribución de los beneficios, con los recursos disponibles y a suponer que su derecho a los últimos implica un derecho correlativo a los primeros, los cambios tecnológicos que ayudan a aumentar la producción agrícola e industrial han de tener efectos adversos en la estabilidad social, pues los agricultores se encontrarían en una posición cada vez menos favorable respecto de los propietarios y los obreros de las fábricas y supondrían que se los estaría privando de sus justas recompensas. Este problema ha constituido un tema permanente en la historia de los Estados Unidos desde fines del siglo XIX.

El economista defensor del bienestar puede tratar de afrontar el problema del crecimiento industrial y la depresión agrícola mediante medidas políticas que inflen artificialmente el valor de los recursos disponibles de la agricultura —subsidios a los precios agrícolas, límites de producción, etc.—. Tales medidas producen una colocación cada vez más errada de los recursos sociales y una ineficiencia general siempre creciente de la producción respecto de la demanda real. Y mientras tanto el sector industrial de la economía se resiente a causa de estas medidas, correctamente consideradas perturbadoras de los efectos de la interacción del mercado, aunque ese resentimiento pueda ser relativamente débil porque sólo implica un promedio menor de ganancia antes que una pérdida real. La única forma de reinstaurar la eficiencia es pues recolocar los recursos, incluido el personal, del terreno agrícola al terreno industrial. Al no existir la interferencia del estado providente, esa recolocación puede darse bajo presión económica, con lo cual la reinstauración de la eficiencia tiene graves costos a corto plazo para las personas que deben modificar su actividad. Y ese costo puede resultarles especialmente desagradable de aceptar, porque con todo derecho, esas personas sienten que se les socava su estilo de vida. Pero el mercado no se caracteriza por respetar los estilos de vida.

A fin de minimizar la ineficiencia de una manera compatible con la maximización del bienestar, es necesario cambiar los recursos disponibles: recolocar los derechos. Pero esto socava el carácter fijo que deben tener los derechos. Una vez que reconozcamos que para mantener la correcta relación utilitarista entre los recursos disponibles y la distribución de los beneficios debemos ajustar continuamente esos recursos a la luz de los cambios tecnológicos, necesariamente llegaremos a la conclusión de que los llamados derechos se desdibujan en una sociedad dinámica utilitarista.

Los cambios producidos en la función de la demanda, que reflejan los cambios de las preferencias individuales, tienen las

mismas consecuencias. Cuando la demanda cambia, ciertos recursos aumentan su provecho y otros lo disminuyen. Es típico que los cambios de la demanda aparezcan acompañados de los cambios de la tecnología: el ejemplo clásico de la demanda de los servicios de los herreros, ilustra bien esa situación. Por lo tanto debemos reemplazar nuestro panorama inicial de dotes fijas determinadas por su maximización del bienestar, por un nuevo panorama en el cual las dotes constituyan sólo una barrera temporal contra la redistribución directa, pero se ajusten a los cambios que se den en la producción en el intercambio o en la demanda.

Este nuevo panorama debería detener el estusiasmo por alabar el utilitarismo. En condiciones estáticas parece aceptable basar los derechos en el principio utilitarista. Pero la aparente plausibilidad que tiene el hecho de aceptar esto se hace aún más aparente cuando analizamos la dinámica de una sociedad en la cual los derechos llegan a estar sujetos a un continuo reajuste. Parece entonces evidente que, para citar una vez más a Winch, "los derechos a los recursos pasan al gobierno y que la asociación de tales derechos con los individuos es una cuestión temporal de conveniencia".[20] Aunque esa cesión de derechos está implícita en el enfoque utilitarista, queda oculta si los cambios sufridos por los derechos como resultado del desarrollo económico se ignoran.

El utilitarismo les niega a los individuos que forman parte de una sociedad la libertad de que gozaba Robinson Crusoe quien podía poner sus habilidades al servicios de sus preferencias. Pues el utilitarista supone que hasta los atributos naturales de una persona, sus habilidades físicas y mentales, le pertenecen sólo en la medida en que esa posesión sea conveniente desde el punto de vista social. Podría afirmarse que el bienestar queda casi con certeza maximizado si cada persona goza de un derecho a aquello de lo que naturalmente puede hacer uso y que ningún otro puede utilizar en su ausencia, es decir, un derecho a aquello que hemos llamado la dote básica de cada persona. Pues puede parecer evidente que el empleo eficiente de esas habilidades se alcanza dándole la plena posesión del derecho a la persona que tiene un acceso natural o directo a esas habilidades. Imaginemos un esquema como el de

la *demoktesis*, esbozado por Robert Nozick, en el cual cada persona pertenece a todas, de modo tal que el derecho al uso de las habilidades físicas y mentales de cada persona se divide y se vende en porciones iguales a cada miembro de la sociedad.[21] En este esquema los costos de tomar decisiones llegarían rápidamente a ser prohibitivos, por lo tanto sería dable esperar que el consorcio que me posee se pusiera de acuerdo para permitirme tomar las decisiones importantes referentes al mejor uso que puedo hacer de mi talento y de mis fuerzas. O por lo menos debería permitirme tomar las decisiones de nivel más bajo. Pero si yo fuera el mejor cardiocirujano del mundo, ese consorcio podría no querer concederme mi derecho ejecutivo de tan buena gana. Deberíamos esperar que el utilitarista dé la bienvenida a la idea, propuesta por John Rawls, de que los talentos de los seres más favorecidos naturalmente constituyen una posesión común para el beneficio de todos.[22]

Si la identificación de cada persona con sus dotes básicas no es arbitraria, como afirmamos en la sección 3.2, la determinación utilitarista de los derechos que cada persona tiene a los recursos basada en la maximización del bienestar le niega pues la actividad libre al individuo. Por consiguiente, rechazamos el intento utilitarista de conciliar la maximización del bienestar con la operación del mercado perfectamente competitivo que establece una forma inaceptable de determinar los recursos disponibles individuales, que en realidad le niega a cada persona su dote básica de habilidades físicas y mentales. Pero el utilitarismo es una teoría proteica y cuando consideremos la naturaleza de las restricciones morales con las que las personas deben estar racionalmente de acuerdo, comprobaremos que nuevamente se opone a nuestra argumentación.

4.2 El marxismo presenta una oposición directa a nuestro enfoque de la operación del mercado que, si pudiera sustentarse, refutaría la afirmación de que la interacción del mercado es imparcial. Pues los marxistas sostienen que la posesión privada de los medios de producción (un supuesto fundamental del mercado) es necesariamente explotadora. El argumento es simple.[23] Si la posesión de los medios de producción está en manos privadas, nada puede evitar que aparezca una situación en la cual algunos individuos (los capitalistas) posean los medios

que otros (los obreros) necesitan si quieren participar de la actividad productiva. Estos otros se ven así obligados a vender sus fuerzas laborales a los propietarios de los medios materiales que exige la producción. Esa venta es explotadora. Porque la característica esencial y distintiva del trabajo es que produzca un beneficio superior a los costos de su propia producción, de modo que el trabajo se reproduce a sí mismo y produce además lo que el pensamiento marxista llamó la *plusvalía*. Ahora bien, la fuerza laboral se compra y se vende, como cualquier otra mercadería, a un precio suficiente para cubrir su costo de producción. De ahí que quien compra la fuerza laboral reciba necesariamente la plusvalía puesto que le paga al trabajador un salario igual al costo de producir la fuerza laboral vendida y recibe un precio igual al valor de lo que la fuerza laboral produce. El mercado favorece sistemáticamente al comprador de la fuerza laboral antes que al vendedor, de ahí que su operación sea en principio parcial, es decir que beneficie al capitalista.

La forma en que los marxistas presentan este argumento lo hace depender de la teoría laboral del valor, que sostiene que el valor de una mercancía está determinado por la cantidad del trabajo necesario desde el punto de vista social exigido para su producción.[24] Esto, por supuesto, es incompatible con nuestra identificación del valor y la utilidad. Pero podemos discutir la afirmación de que la plusvalía va para el comprador antes que para el vendedor de la fuerza laboral sin refutar primero la teoría marxista del valor. Porque los marxistas en su conjunto están de acuerdo en afirmar que, sea cual fuere el valor de una mercadería, su precio de mercado está determinado por la oferta y la demanda. De modo que sencillamente podemos preguntar si el vendedor de la fuerza laboral resulta perjudicado por los términos de la venta, si el precio que recibe por su fuerza laboral le proporciona una ventaja sistemática al comprador.

Comencemos por la afirmación de que el precio del mercado de aquello que produce el obrero es mayor que el costo de su propia producción. Supongamos que la fuerza laboral se vende a un precio igual al costo de producirla. Entonces, evidentemente, habrá una demanda de más trabajo, puesto que los compradores se benefician con la diferencia que existe entre los que pagan por la fuerza laboral y lo que reciben por su producto. Esa

demanda continuará hasta que el precio recibido por el producto marginal de una unidad adicional de fuerza laboral sea igual al costo marginal de producir esa unidad adicional. Pero al llegar a ese punto, el salario —el precio pagado por la fuerza laboral— sería igual al precio recibido por su producto. Cuando se alcanza un equilibrio entre la oferta y la demanda de trabajo ya no hay ninguna plusvalía. Los trabajadores reciben un salario igual a la diferencia marginal que su trabajo agrega al producto total.

Los marxistas evitan esta conclusión al negar que la oferta y la demanda puedan estar en equilibrio. E insisten en afirmar que el comprador de la fuerza laboral puede mantener el salario por debajo del precio que recibe por su producto, porque la oferta de trabajo excede su demanda, puesto que siempre hay "un ejército de reserva de desempleados".[25] Pero acabamos de mostrar que si el salario está por debajo del precio recibido por el producto del trabajo, debe de haber pues una demanda efectiva de más trabajo, es decir que ha de darse una demanda que exceda la oferta. Por lo tanto el marxista queda atrapado en una contradicción. El comprador de la fuerza laboral puede obtener una plusvalía —pagar un salario inferior al precio que recibe por el producto del trabajo— sólo si la oferta de trabajo excede la demanda. Pero si hay una plusvalía por obtener, esto crea una demanda de trabajo que excede la oferta existente. O, para decirlo de otro modo, si la oferta de trabajo excede su demanda, ello sólo puede deberse a que el costo de producir el trabajo excede el precio que puede obtenerse por su producto. De modo que sólo puede haber plusvalía si la oferta excede la demanda, pero si la oferta excede la demanda no puede haber ninguna plusvalía.

La operación del mercado perfectamente competitivo debe producir un equilibrio entre la oferta y la demanda de fuerza laboral, si —como suponen los marxistas— es una mercancía creada por el mercado. Si existe el equilibrio, probablemente el comprador no tenga ninguna plusvalía que obtener y por lo tanto tampoco habría explotación del vendedor. Ahora bien, esto no prueba nada en cuanto a la operación de los mercados imperfectamente competitivos. Tampoco prueba nada de una situación en la cual la fuerza laboral aparece como un subproducto de otras actividades de las cuales la gente participa por

razones completamente distintas de adaptarse a la demanda laboral del mercado. Esto podría dar lugar a un exceso permanente de fuerza laboral de la que se beneficiaría el propietario de los medios de producción. Y esto podría plantear interrogantes referentes a la propiedad de los recursos disponibles que colocan a ciertas personas en posición de beneficiarse gracias a esa sobreoferta de trabajo.

La aparición de la explotación del mercado puede haberse debido a la explotación real misma ocasionada en verdad por los rasgos de no mercado de la sociedad, o por rasgos, tales como la distribución inicial de los recursos, considerada como algo dado en la interacción del mercado. Y esto puede ayudar a explicar por qué el argumento marxista ha convencido a tanta gente, aun cuando en realidad sea incoherente y esté erróneamente dirigido. Pero si es un argumento erróneamente dirigido, que no tiene nada que ver con la operación de los mercados perfectamente competitivos, la teoría marxista de la plusvalía y la explotación sólo sirve como una pantalla de humo que oculta la verdadera explotación que pueden encarnar los arreglos sociales reales, incluyendo, por supuesto, aquellos de las sociedades "marxistas".

Notas

1 Adam Smith, *An Enquiry into the Nature and Causes of the Wealth of Nations*, 1776, párrafo iv, cap. 9, presentado por E. R. A. Seligman, 2 vols., Londres, 1910, vol.2, pág. 180.

2 *Ibíd*, párrafo iv, cap. 2, vol 1, pág. 400.

3 Hobbes, *Leviathan*, cap. 13, pág. 63.

4 Véase J. C. Harsanyi, "Advances in Understanding Rational Behavior", en *Essays on Ethics, Social Behavior, and Scientific Explanation*, Dordrecht, 1976, pág. 94.

5 J. Rawls, *A Theory of Justice*, Cambridge, Mass. 1971, pág. 13.

6 Véase P. H. Wicksteed, *The Common Sense of Political Economy and Selected Papers and Reviews on Economic Theory*, L. Robbins, 2 vols., Londres, 1933; "El carácter específico de una relación económica es no su 'egoísmo' sino su 'no-tuismo'", vol.1, pág. 180.

7 Véase Karl Marx, *The Capital: A Critical Analysis of Capitalist Production*, vol i, cap. 1, sec. 4; trad. S. Moore y E. Aveling, Londres, 1889, págs. 41-55.

8 Sobre una breve explicación de esta demostración que pone el acento en que la competencia perfecta es suficiente pero no necesaria para la optimidad, véase D. M. Winch, *Analytical Welfare Economics*, Harmondsworth, Middx., 1971, págs. 89-95.

9 Véase Wicksteed; "las fuerzas económicas...tienden a asegurarle a todo tipo de esfuerzo humano una remuneración correspondiente al valor maginal que tiene ese esfuerzo para cualquier miembro de la comunidad.", vol. 1, pág. 338.

10 Winch, *op. cit.*, pág. 97.

11 Sobre una breve análisis del rédito económico en relación con los rendimientos del mercado, véase S. Gordon, *Welfare, Justice and Freedom*, Nueva York, 1980, págs. 98-99.

12 Hobbes, *Leviathan*, cap. 14, pág. 64.

13 Véase Immanuel Kant, *Groundwork of the Metaphysic of Morals*, 1785, en *Kants Gesammelte Schriften*, 22 vols., Berlín, 1900-1942, vol. 4, págs. 397-398, 444.

14 Véase Gayle Rubin, "The Traffic in Women: Notes on the 'Political Economy' of Sex", en Rayna R. Reiter (comp.), *Toward an Anthropology of Women*, Nueva York, 1975, especialmente las págs. 173-177. Como hace notar Rubir, la frase "tráfico de mujeres" procede de Emma Goldman.

15 Véase Hume, *Treatise*, ii, i, xi, págs 316-320 y iii, i, págs. 575-578.

16 Véase, además de la referencia de la nota 13, Kant, *Critique of Practical Reason*, en *Kants Gesammelte Schriften*, vol.5, págs. 22-26.

17 Winch, *op. cit.*, pág. 97.

18 Mill, *op. cit.*, cap. 5, párrafo 25.

19 Véase Winch, *op. cit.*, pág.193.

20 *Ibid.*, pág. 99.

21 Véase R. Nozick, *Anarchy, State, and Utopia*, Nueva York, 1974, págs. 280-290,

22 Véase Rawls, *A Theory of Justice*, págs 101-102.

23 Véase Marx, *The Capital*, vol. 1, cap. VII, sec. 2, págs. 166-180.

24 *Ibid.*, vol. 1, cap. I, sec.1, pág. 6.

25 Frederick Engels, "The Housing Question", en Karl Marx y Frederick Engels, *Selected Works*, 2 vols., Moscú, 1951, vol.1, pág. 527. La expresión usual es "ejército de reserva industrial", como en Marx, *The Capital*, vol 1, cap. XXV, sec. 3, pág.642.

V

Cooperación: negociación y justicia

1.1 La razón, que aumenta los costos de la interacción natural entre los seres humanos, ofrece no solamente un remedio para todos los males que crea, sino también la perspectiva de nuevos beneficios generales que han de alcanzarse mediante la cooperación. Cuando la mano invisible no logra hacer que cada persona, atenta únicamente a su propia ganancia, tienda a promover el beneficios de todos, la cooperación llega a ser una mano visible. En este capítulo veremos que la cooperación constituye una respuesta racional a las fallas del mercado.

Cuando la interacción del mercado, con su preestablecida armonía entre lo equilibrado y lo óptimo, está más allá del bien y del mal y cuando la interacción, con la aparición de los jugadores líberos y los parásitos, degenera en una situación en la que tallan la fuerza y el fraude, la interacción cooperativa es el terreno de la justicia. La justicia es la disposición que tiene el individuo a no sacar ventaja de sus semejantes, a no tratar de apoderarse de los bienes libres y a no imponer costos no compensados siempre que ese individuo suponga que los demás están igualmente dispuestos. Mostraremos que al satisfacer las condiciones de racionalidad práctica, la cooperación garantiza la desaparición de la posición de jugador líbero y de parásito endémica en nuestra condición natural, de modo que podemos identificar la justicia con la disposición racional a adoptar una conducta cooperativa. Por consiguiente coincidimos con el más influyente teórico de la justicia contemporáneo, John Rawls, cuando dice: "Podemos caracterizar las circunstancias de la justicia como las condiciones normales en las cuales la cooperación humana es posible y necesaria".[1]

David Hume, en cuyas ideas se basa Rawls al hablar de la circunstancias de la justicia, sostiene que la necesidad de la cooperación surge de la conjunción de la escasez, que caracteriza nuestras "circunstancias exteriores", y una desviación en favor del sí mismo, que caracteriza nuestro "temperamento natural".[2]

El desinterés mutuo supuesto por el mercado es una forma extrema de esa desviación hacia uno mismo, aunque la estructura de la interacción del mercado la considere inocua. Si la naturaleza proporcionara en abundancia los bienes necesarios para satisfacer nuestros deseos, o si la benevolencia hiciera que cada persona considerara los intereses de sus semejantes como los propios, no habría jugadores líberos ni parásitos, cuyas ambiciones deban quedar restringidas por la mano visible de la cooperación. Todos tratarían naturalmente de coordinar sus acciones a fin de lograr el bien común, sin presentar reclamos opuestos del fruto de sus esfuerzos que la justicia está obligada a resolver.

Pero la escasez y la desviación hacia uno mismo no constituyen una justificación suficiente para que se dé la interacción cooperativa. Si la escasez que afronta cada persona no estuviera agravada por la presencia de los demás, por más desviada hacia sí misma que esté esa persona, las actividades que desarrolle tendrían poca relación con las de los demás, de modo que no habría conflicto ni cooperación. Y esa escasez tampoco podría aliviarse mediante las actividades conjuntas, por lo tanto el terreno de la justicia se extendería únicamente a evitar los conflictos mutuamente destructivos y no a proveer cooperativamente beneficios mutuos. Por consiguiente, debemos incluir entre las circunstancias necesarias para la justicia, la variabilidad de la escasez. Las fuentes de la satisfacción y la no satisfacción no constituyen una oferta fija, de modo que una interacción apropiada debe disminuir los costos generales y aumentar los beneficios generales.

Ahora bien, estrictamente hablando, lo importante no es el hecho de la oferta variable, sino la conciencia de ese hecho. Esa conciencia tiene tanto efectos positivos como efectos negativos. Por un lado, tomamos conciencia de que cualquier otro es un competidor que tiende a obtener esos bienes escasos y esa conciencia exacerba nuestra competencia y aumenta nuestros costos. A esa conciencia nos referimos cuando decimos que la razón se suma a las desventajas de la naturaleza. Por otro lado, cobramos conciencia de que cada uno de los demás es una persona que potencialmente ha de cooperar en la producción de una mayor oferta de bienes, y esa conciencia nos permite

advertir la posibilidad de nuevos beneficios. Por lo tanto observemos también que la razón ofrece ventajas que no pueden obtenerse en el estado natural.

Los diferentes teóricos varían en cuanto a la importancia que le dan a esos aspectos de nuestra conciencia. Tomas Hobbes pone el acento en el efecto que causa reconocer a los demás como competidores.[3] Y sostiene que, teniendo en cuenta la escasez y el desinterés mutuo, cada individuo debe considerar a todos los demás por lo menos en competencia por los bienes que él necesita para subsistir o para obtener un bienestar mayor. Pero esto crea en cada persona una preferencia real por dominar a sus semejantes; si esa persona es capaz de establecer tal dominio ahora, puede pues esperar tener más éxito en cualquier disputa futura por los bienes escasos. Así el conflicto potencial se transforma en hostilidad real que conduce a todos los hombres a una guerra contra todos los demás, una guerra de la que cada uno debe esperar sufrimiento, aun cuando nadie se arriesgaría a abstenerse de ella. A fin de impedir esa guerra interminable cada individuo debe aceptar restringir su conducta, siempre que los demás también lo acepten, con el objetivo de que todos puedan vivir pacíficamente. Para Hobbes, la justicia se basa en la aceptación de tal acuerdo.

El enfoque de Hobbes pone claramente de manifiesto la función que cumple la razón en el aumento de los costos de la interacción natural. Para Hobbes un ser razonable que busca maximizar su utilidad es aquel que adopta una estrategia que le asegure el derecho de prioridad en su interacción con los demás. Aquellos que no cobran conciencia de la perspectiva de beneficiarse imponiéndoles costos a los demás, no han de encontrarse en la condición particularmente desagradable que describe Hobbes. Pero la explicación que da Hobbes revela igualmente la función que cumple la razón en cuanto a permitirnos superar esa condición, porque la razón indica los "convenientes Artículos de la Paz" y quienes se adhieran a ellos pueden salir de la condición natural de la humanidad y finalmente entrar en la sociedad.[4]

En la argumentación que desarrolla Hobbes hay muy poco que sugiera una función más positiva de la cooperación. Somos conscientes de que cada uno de los demás es un competidor

entonces estamos dispuestos a cooperar a fin de evitar el conflicto mutuamente destructivo, pero somos menos conscientes de que cada uno de los demás es una fuente potencial de beneficio mutuo. En realidad, el mundo que describe Hobbes se parece en algunos aspectos al que, según se dice (no sé con cuánto conocimiento de causa) se les ofrece a los habitantes de las islas del Pacífico Sur, los dobus.[5] Como lo que nos importa es la idea y no la realidad, podemos ignorar los datos antropológicos, que pueden confirmar o desmentir nuestro relato, y suponer que los dobus creen que el mundo les ofrece una cantidad —oferta— fija de los bienes que ellos atesoran, principalmente batatas. Cuantas más batatas tenga en mi huerta, menos tendrás en la tuya. En el esquema de los dobus no hay lugar para ideas tales como la cooperación dirigida a aumentar los beneficios, puesto que no pueden crecer más batatas. A lo sumo, pueden evitarse o disminuirse los costos del conflicto hobbesiano. Pues aun cuando cada persona se sintiera inclinada a adquirir, mediante los apropiados artificios mágicos, las batatas de la huerta de su vecino, seguiría siendo mejor para cada uno si ninguno tratara de imponerse a los demás mediante artificios mágicos o de cualquier otro tipo.

El mundo de los dobus puede tener sus rebuscados defensores. En realidad, si consideramos la interacción entre las clases sociales antes que la que se desarrolla entre los individuos, los marxistas estarían entre tales defensores, pues suponen que en una sociedad basada en la división de clases, la mayor parte de las batatas pertenece a la burguesía y la menor cantidad corresponde al proletariado. Y esto lleva a algunos marxistas, insuficientemente atentos a Hobbes, a negar la posibilidad de la cooperación y la importancia que tiene la justicia en las sociedades basadas en la división de clases. Pero nosotros suponemos que la visión del mundo que tienen los dobus es falsa. Y, teniendo en cuenta que la oferta es variable, podemos ser conscientes de que los demás son individuos que potencialmente han de incrementar la producción. Esta es la conciencia sobre la que ponen el acento Hume y Rawls.

Pero el hecho de reconocer las posibilidades de incrementar la producción que están implícitas en la idea de la oferta variable es también la base del mercado. Y tanto Hume como Rawls

pasaron por alto el papel que desempeña el mercado en limitar la necesidad de la cooperación. La interacción del mercado da condiciones de oferta variable entre personas que sienten mutuo desinterés. Pero si se da la competencia perfecta, esas personas no tienen necesidad de una interacción cooperativa que constituya la mano visible, puesto que la optimidad del resultado del mercado excluye cualquier alternativa que podría reducir los costos generales o aumentar los beneficios generales. Por lo tanto debemos incluir entre las circunstancias de la justicia, además de la conciencia de la escasez variable y la desviación individual, el reconocimiento de la presencia de las externalidades. La justicia debe corregir tanto las ineficiencias ocasionadas por la imposición de costos que no proporcionan los correspondientes beneficios, como aquellas ocasionadas cuando no se pueden proporcionar beneficios debido a la incapacidad de cubrir los costos. Puesto que las externalidades suponen la existencia de una oferta variable, podemos pues decir que las circunstancias fundamentales de la justicia , es decir aquellos rasgos de la situación humana que originan la cooperación, son la conciencia de que hay externalidades en el ambiente que nos rodea y la conciencia de que nuestro carácter tiende a desviarse hacia uno mismo.

1.2 Cuando el mercado fracasa, cada persona que busca maximizar su utilidad atendiendo a las estrategias que espera que los demás elijan, no puede maximizar su utilidad atendiendo a las utilidades que reciben los demás. El resultado en equilibrio de las respuestas mutuamente maximizadoras de la utilidad no es óptimo. El dilema de la racionalidad estratégica, acabadamente planteado en el dilema del prisionero, llega a apoderarse de nuestras interacciones y a alejarlas de los objetivos a los que apuntamos. Hasta ahora hemos supuesto que cada persona se forma expectativas referentes a las elecciones de los demás a las cuales responde eligiendo su propia estrategia, pero que esa persona sólo considera los costos y los beneficios que le implica tomar esa decisión. Es decir que hemos ignorado las externalidades, un tercer factor no elegido. De modo que en muchas situaciones, el resultado al que se llega mediante las elecciones independientes de aquellos que participan de la interacción es subóptimo. A fin de tener efectivamente

en cuenta las externalidades, cada persona debe elegir su estrategia para llegar a un resultado particular determinado por un acuerdo previo establecido entre los participantes. Ese acuerdo, si es racional, ha de asegurar la optimidad. Por supuesto, ese acuerdo puede estar implícito antes que explícito, puede ser una convención o un entendimiento general antes que un contrato. Pero no es una mera ficción, puesto que da lugar a un nuevo modo de interacción que identificamos con el nombre de cooperación. En la interacción natural cada persona afronta un problema estratégico separado que debe resolver mediante la elección de su estrategia. En su carácter de cooperadoras, todas las personas afrontan un problema estratégico común que debe ser resuelto para poder determinar la elección estratégica de cada individuo.

Como una caracterización preliminar, que luego modificaremos en la próxima subsección, digamos que en una interacción cooperativa, el objeto primario de elección es un conjunto de estrategias, una correspondiente a cada persona. Si la cooperación es voluntaria, como supondremos salvo en los casos en que establezcamos explícitamente lo contrario, cada individuo participa de esa elección primaria o está de acuerdo con ella. Puesto que el producto de un conjunto de estrategias —cada una correspondiente a cada persona que participa de la interacción— es un resultado, podemos decir también que en la interacción cooperativa el objeto del acuerdo es un resultado, el cual determina la elección estratégica de cada persona. En este capítulo estableceremos las condiciones para una elección cooperativa racional o para un acuerdo racional sobre un resultado.

Y hemos establecido una de esas condiciones: el objeto de una elección cooperativa racional debe ser un resultado óptimo. En el capítulo 3 observamos que un resultado puede considerarse o bien un producto de estrategias, o bien un conjunto de utilidades. Deberíamos esperar que si cada persona selecciona su estrategia independientemente de los demás, como ocurre en la interacción natural y en el mercado, la primera de esas concepciones sería la que prevalece. Pero si cada individuo elige su estrategia como el resultado de un acuerdo con los demás, puesto que cada uno al alcanzar el acuerdo está interesado en su propia utilidad, y no en su estrategia, deberíamos esperar

que la segunda concepción fuera la preeminente. Ahora bien, esto indica que en la interacción no cooperativa la propiedad fundamental de la racionalidad es el equilibrio, mientras que en la interacción cooperativa la propiedad fundamental de la racionalidad es la optimidad. Por consiguiente es necesario interpretar las tres condiciones presentadas en el capítulo 3, de un modo cuando se trata de una interacción no cooperativa natural y de mercado, y de otro modo cuando se trata de una interacción cooperativa. La condición *A* —que sostiene que la elección de cada persona debe constituir una respuesta racional a las elecciones que esa persona espera que los demás hagan— debe interpretarse como una condición que exige una respuesta maximizadora de la utilidad en la interacción no cooperativa y que exige una respuesta optimizadora en la interacción cooperativa. O, en otras palabras, cuando no existe un acuerdo sobre un resultado o sobre un conjunto de estrategias, es racional que cada persona intente maximizar su utilidad atendiendo a las estrategias que esa persona espera que elijan los demás, mientras que en el contexto de un acuerdo, es racional que cada persona apunte a obtener un resultado óptimo atendiendo a las estrategias acordadas de los demás. La primera parte de esta proposición es un corolario directo de identificar la racionalidad práctica con la maximización de la utilidad individual, pero el corolario puede parecer tan directo que termina por refutar la segunda parte de la proposición. Nuestra tarea consiste pues en mostrar que no es así, que la cooperación es una forma racional de interacción.

En los dos párrafos anteriores planteamos dos problemas por completo diferentes. Supongamos que tenemos que ponernos de acuerdo respecto de un resultado o de un conjunto de estrategias; ¿en qué condiciones podríamos decir que nuestro acuerdo es racional? Este es el primer problema y se refiere a lo que podemos llamar la racionalidad interna de la cooperación, la forma racional de tomar una decisión cooperativa. Podríamos resolver el problema si estableciéramos un principio que racionalice el acuerdo del mismo modo en que el principio de la maximización de la utilidad esperada racionaliza la elección individual.

¿En qué condiciones es racional llegar a un acuerdo respecto de un resultado o de un conjunto de estrategias y obrar ateniéndonos a ese acuerdo? Este es el segundo problema y se refiere a la racionalidad externa de la cooperación. Debemos mostrar cuándo es racional obrar de manera cooperativa en lugar de hacerlo de manera no cooperativa. Al resolver el primer problema damos por sentada la cooperación; al afrontar el segundo problema ponemos en tela de juicio la cooperación y exigimos su racionalidad como un modo de interacción. Por supuesto, ambos problemas están relacionados entre sí, pero en este capítulo nos ocuparemos únicamente del segundo en la medida en que éste lleva implícito el primero. Y lo resolveremos en los capítulos 6 y 7.

Hay un tercer problema que se refiere a la moralidad de la cooperación. ¿En qué condiciones la interacción cooperativa es justa o imparcial? En este caso, como en el caso del mercado, debemos distinguir la operación de las prácticas cooperativas de sus condiciones. En este capítulo examinaremos la imparcialidad de la operación de las prácticas y las instituciones cooperativas a la luz de nuestro enfoque de la racionalidad interna de la cooperación, preguntándonos si el principio de elección cooperativa racional es un principio imparcial. Como hay un significativo paralelismo entre la argumentación desarrollada en este capítulo y la desarrollada en el capítulo anterior, debemos poner de manifiesto la racionalidad y la imparcialidad de la interacción cooperativa, exactamente como pusimos de manifiesto la racionalidad y la imparcialidad de la interacción del mercado. Pero comprobaremos que la cooperación no crea una zona moralmente libre y que, en cambio, exige que se apliquen restricciones morales a la conducta maximizadora.

En la próxima subsección echaremos luz sobre las características de la interacción cooperativa al mostrar cómo ésta puede aumentar la cantidad de estrategias disponibles y de los resultados (esperados). En la sección 2 analizamos los distintos enfoques que permiten determinar las condiciones que debe satisfacer la elección cooperativa para ser racional. En la sección 3 formulamos detalladamente nuestra propia respuesta; desarrollamos un principio para el acuerdo racional. Y en la sección 4 examinamos la imparcialidad de la cooperación racional.

1.3 Ernie y Bert quieren encontrarse. Ernie preferiría

hacerlo en la biblioteca y Bert en el cine. Y cada uno iría indistintamente a cualquiera de los dos sitios si se encontrara con su amigo. De modo que podríamos establecer una medida de intervalo de las preferencias de ambos. Supongamos que Ernie es indiferente entre encontrarse en el cine y una lotería que dé idénticas probabilidades a encontrarse en la biblioteca y a no encontrarse y que Bert sea indiferente entre encontrarse en la biblioteca y una lotería que le asigne las mismas probabilidades a encontrarse en el cine que a no encontrarse. Al calcular las utilidades de ambos amigos de acuerdo con el método explicado en 2.3.2 podemos representar la situación mediante el siguiente esquema:

	Bert va a	
	la biblioteca	el cine
Ernie va a la biblioteca	1, 1/2	0, 0
Ernie va al cine	0, 0	1/2, 1

Hay exactamente dos resultados óptimos, ambos amigos deben ir al mismo lugar. Esos resultados están además en equilibrio. Desafortunadamente, la simetría de la situación hace imposible que Ernie y Bert coordinen sus acciones a fin de obtener uno de esos resultados utilizando el principio de Zeuthen, como sí podían hacerlo Victor y Valerie en 3.2.2. Si cada uno elige ir al lugar que prefiere para encontrarse corre el riesgo de que el encuentro no se realice. Si cada uno, indeciso, opta por echar suerte con una moneda (es decir adopta una estrategia mixta que le asigna una probabilidad de 1/2 a cada una de las acciones posibles), el resultado esperado ha de ser pues una lotería con una probabilidad de 1/4 de que se encuentren en la biblioteca, una probabilidad de 1/4 de que se encuentren en el cine y una probabilidad de 1/2 de que el encuentro no se produzca; y la utilidad esperada ha de ser para cada amigo de 3/8.

Pero supongamos que Ernie y Bert adoptan un enfoque cooperativo de la interacción. En ese caso se abre toda una nueva gama de posibilidades. Porque los amigos pueden hacer una única elección estratégica conjunta antes que elecciones estratégicas independientes. Y la simetría de la situación que los incluye señala una única estrategia conjunta que le asigna una probabilidad de 1/2 al hecho de que ambos vayan a la

biblioteca y una probabilidad de 1/2 al hecho de que ambos vayan al cine. En lugar de que cada uno eche al aire una moneda, ambos tiran una sola moneda y aceptan obrar de acuerdo con el resultado. El resultado esperado es pues una lotería que le asigna una probabilidad de 1/2 al hecho de que se encuentren en la biblioteca y una probabilidad de 1/2 al hecho de que se encuentren en el cine; ambos tienen así una utilidad esperada de 3/4. Este resultado es óptimo como en realidad lo es el resultado de cualquier estrategia mixta conjunta que consiste en una lotería que tiene únicamente por premios que ambos vayan a la biblioteca y que ambos vayan al cine. Pero, en una interacción no cooperativa Ernie y Bert no disponen de ninguna de esas estrategias y por lo tanto de ninguno de esos resultados.

Como lo ilustra este ejemplo, la cooperación puede ampliar la cantidad de estrategias y de resultados (pero no siempre lo hace; Noreen y Norman [3.2.2] no pueden ampliar sus perspectivas mediante la cooperación). Definamos una *estrategia conjunta pura* como el producto de los elementos de un conjunto de estrategias puras, cada una correspondiente a cada persona que participa en la interacción. Una estrategia conjunta pura es pues simplemente un resultado posible (si excluimos los resultados meramente esperados). Una *estrategia conjunta mixta* es pues una lotería que tiene por premios estrategias conjuntas puras o resultados posibles. Un resultado esperado es pues simplemente una estrategia conjunta mixta o una lotería sobre los posibles resultados. En la interacción cooperativa cada lotería sobre los resultados posibles determina un resultado esperado. Pero no necesariamente es ése el caso en la interacción no cooperativa, en la cual el conjunto de resultados esperados es un subconjunto (a veces adecuado y a veces inadecuado) del conjunto de loterías sobre los posibles resultados.

Al extender el alcance de las estrategias de lo individual a lo conjunto, alteramos implícitamente la caracterización previa que habíamos hecho de la interacción cooperativa, según la cual ésta tenía como objeto del acuerdo un conjunto de estrategias, una correspondiente a cada persona. Ahora tendríamos que decir que el objeto primario de la elección o el acuerdo es una estrategia conjunta. Puesto que un conjunto de estrategias, una correspondiente a cada persona, puede quedar representada

por una estrategia conjunta y puesto que no es posible lo contrario, esta nueva caracterización amplía nuestra concepción inicial de la cooperación.

Cualquier interacción cooperativa en la que participen sólo dos personas puede representarse convenientemente en forma gráfica. Supongamos que el eje horizontal representa la escala de utilidad de una persona y que el eje vertical representa la escala de utilidad de otra persona. Cada resultado posible queda representado pues si se consideran las utilidades como coordenadas. La figura convexa cerrada que se obtiene al unir los puntos que representan los resultados posibles representa pues la cantidad de posibles loterías sobre los resultados posibles o sobre los resultados esperados. Cada punto que se sitúa dentro de esta figura o sobre ella corresponde a un resultado esperado y cada resultado esperado corresponde a uno de esos puntos. Llamamos a esta figura el *espacio resultado.* El límite superior derecho de este espacio representa pues los resultados óptimos. Las figuras 1 y 2 ilustran dos de nuestros ejemplos —Ernie y Bert y Jane y Brian del capítulo 3—. En la figura 2 mostramos también el punto que representa el único resultado en equilibrio que evidentemente no es óptimo.

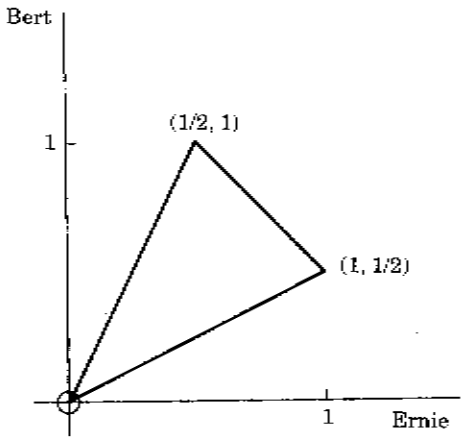


Figura 1

Aunque el tratamiento gráfico llegue a ser dificultoso, esta forma de representación puede extenderse a las interacciones cooperativas de las que participen cualquier cantidad de personas, utilizando un espacio de utilidad dimensional *n,* en el que a cada dimensión correspondan las utilidades de cada persona. El espacio-resultado continúa siendo una figura convexa cerrada aun cuando se hable de n dimensiones y no de dos y su límite superior, generalmente *n*-1 dimensional, representa los resultados óptimos.

Aunque no pretendamos darle a nuestro enfoque un tratamiento matemático, esbozaremos aquí el problema formal que trataremos de resolver informalmente en este capítulo. Queremos definir una función de elección que llamamos la *función de cooperación.* El dominio de esa función es un conjunto de espacios-resultados. Sus valores son los elementos de un conjunto de puntos del espacio de utilidad. Por supuesto, la idea es que para un espacio-resultado, que representa los resultados de una posible interacción cooperativa, la función determina un punto que representa el resultado particular determinado por la estrategia conjunta de quienes participan de la cooperación.

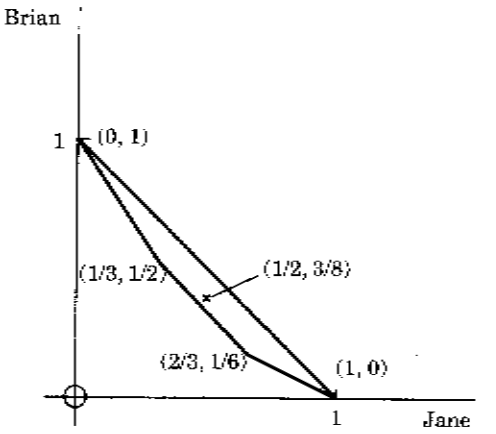


Figura 2

Una condición que trataremos de imponerle a la función cooperativa es que su dominio sea completo, que incluya todos los espacios-resultados posibles. Una segunda condición es que en cada miembro de su dominio la función determine un único punto considerado su valor. Podemos especificar aun más esta condición: el punto debe estar incluido en el espacio-resultado (es decir que representa un resultado de la interacción) y en realidad debe caer sobre el límite superior del espacio-resultado (es decir que representa un resultado óptimo). Estas son sólo condiciones necesarias, demasiado débiles para determinar una función única. Y nuestro objetivo es establecer en la sección 3, aunque sólo sea informalmente, un conjunto de condiciones suficientemente fuertes. Pero comenzaremos considerando tres maneras de llegar a una función de cooperación.

2.1 Varias personas se ponen de acuerdo para perseguir un resultado que por lo tanto debe alcanzarse mediante una estrategia conjunta que determine la acción que ha de realizar cada persona. ¿En qué condiciones debe realizarse ese acuerdo para que sea racional? La elección individual es racional en tanto sea maximizadora de la utilidad. ¿Hay algo análogo a la maximización de la utilidad en el acuerdo en la elección cooperativa? Puesto que el acuerdo ha de ser voluntario, debe reflejar, y en cierto sentido reflejar de manera equitativa, las preferencias de cada persona. Por lo tanto una primera proposición sería que hiciéramos derivar del orden de preferencia individual de quienes participan de la interacción, un orden de preferencias social. Definimos entonces una medida de la preferencia social e identificamos la elección cooperativa racional con la maximización de esa medida.

El lector familiarizado con la bibliografía referente a la elección social reconocerá el problema que afronta esta proposición. Kenneth Arrow demostró que no es posible hacer derivar un orden de preferencia social de cada configuración de los órdenes de preferencia individuales sin violar una de las siguientes condiciones.[6]

(1) Si cada individuo prefiere el resultado X al resultado Y, la sociedad prefiere pues X a Y;
(2) No hay ningún individuo cuyas preferencias sean automáticamente las preferencias de la sociedad, sean cuales fueren las preferencias de los demás;
(3) Las preferencias sociales sobre cualquier resultado dependen sólo de las preferencias individuales sobre tal resultado.

Puesto que necesitamos que la elección cooperativa seleccione un resultado óptimo, no podemos rechazar la condición (1). Puesto que necesitamos que la elección cooperativa refleje las preferencias de cada individuo, no podemos rechazar la condición (2). Y puesto que en esta proposición la preferencia social se basa en las preferencias individuales y sólo en ellas, no podemos rechazar la condición (3). De ahí que no podamos hacer derivar un orden de preferencias social de todas las configuraciones de los órdenes de preferencias individuales, y que la primera sugestión pueda no parecer un buen arranque.

Pero no nos precipitemos demasiado. Suponemos que puede aplicarse una medida de intervalo, la utilidad, a las preferencias de cada cooperador. De modo que una proposición modificada sería que hagamos derivar de esas medidas de intervalo una medida de intervalo social e identifiquemos la elección cooperativa racional con la maximización de esa medida. Sin embargo, esa proposición modificada no nos permite rehuir de la demostración de Arrow. Recordemos que no hemos establecido ninguna base para comparar las utilidades de las diferentes personas. Teniendo en cuenta que es imposible comparar las utilidades interpersonales, la demostración de Arrow puede también aplicarse al problema de definir una medida de intervalo social sobre la base de las medidas de intervalo individuales.[7]

La enunciación formal de nuestro problema presentada en la sección anterior sugiere una modificación más promisoria. Buscamos una función que determine un resultado para todos los espacios-resultados o, en realidad, un resultado para todos los conjuntos de resultados, sobre la base de las utilidades individuales asignadas a los elementos del conjunto. ¿Por qué no pasar por alto la idea de establecer un orden de preferencias social o una medida de intervalo social e ir directamente de las medidas individuales al resultado elegido? Esto puede parecer una evasiva, sin embargo es una evasiva conveniente; podemos hacer derivar una elección social de todas las configuraciones de

los órdenes de preferencia individuales y satisfacer estas condiciones:

(1') Si cada individuo prefiere el resultado *X* al resultado *Y*, la sociedad no ha de elegir *Y* si *X* está disponible;
(2') No hay ningún individuo cuya utilidad quede automáticamente maximizada por las elecciones de la sociedad, sean cuales fueren las utilidades de los demás;
(3') La elección social sobre cualquier resultado depende sólo de las utilidades individuales que proporcione ese resultado.

La proposición revisada es pues que podemos hacer derivar de los órdenes de preferencias individuales de los cooperadores una elección social mediante una función que incluya todas las configuraciones de los órdenes individuales en su dominio y que satisfaga las condiciones (1'), (2') y (3'). Pero la única forma plausible de satisfacer nuestras condiciones es incluir, en la elección social, toda la línea de resultados óptimos.[8] En lugar de escoger un resultado único, comprobamos que no tenemos una base para elegir entre aquellos resultados que satisficieron nuestras condiciones necesarias previas sobre la elección cooperativa.

Pero, una vez más estamos obrando demasiado precipitadamente. En la bibliografía referente a la elección social, generalmente se ignoran las loterías; se considera que los resultados son resultados posibles y se excluyen los resultados esperados que son loterías que tienen como premios los resultados posibles. Puesto que nosotros no tenemos nada que objetarles a las loterías, ¿por qué no podemos suponer que si nuestra función para determinar la elección social selecciona todos los resultados óptimos, esa función selecciona operacionalmente la lotería que le asigna la misma probabilidad a todos los posibles resultados óptimos y que debe ser ella misma un resultado esperado? Esta nueva proposición plantea un problema inmediato; el resultado esperado determinado por una lotería que le asigna la misma probabilidad a cada uno de los resultados óptimos posibles no necesariamente debe ser ella misma óptima.[9] Quizás una nueva modificación pueda solucionar este problema. Pero, puesto que lo que nos interesa no es examinar las cuestiones formales de la teoría de la elección racional, no es

necesario continuar con este análisis. Hemos mostrado que la famosa demostración de Arrow de que es imposible hacer derivar un orden de preferencia social de los órdenes de preferencia individuales; por lo tanto parece factible relacionar la elección cooperativa con las preferencias individuales, de modo tal que la elección en cualquier situación dependa solamente de las utilidades individuales que brinden los resultados de la situación.

Pero por más factible que pueda ser esta relación, nosotros rechazamos la idea de que el acuerdo o la elección cooperativa deban basarse únicamente en las utilidades individuales, pues esa idea lleva a considerar todos los resultados óptimos como resultados indiferentes desde el punto de vista racional. Esta idea puede ser conveniente si pensamos que una elección social es una elección hecha por alguna persona con autoridad o por algún cuerpo que tiene autoridad y que decide teniendo en cuenta las preferencias de las personas afectadas por tal elección y excluyendo cualquier otra consideración. Es decir que se considera a las personas afectadas como seres pasivos. Ninguna interacción que se desarrolle entre ellos puede afectar el resultado. Esas personas son receptoras o consumidoras de los bienes (y quizá de los males) que hay para distribuir; no son productoras de dichos bienes.

El contexto de la elección cooperativa es completamente diferente. Aunque los cooperadores están más interesados en los rendimientos de la utilidad que en los aportes estratégicos de los resultados entre los cuales deben elegir, no por ello son seres pasivos dentro del proceso del acuerdo, un proceso cuyo carácter estratégico no podemos ignorar en este análisis. Cuando existe una cantidad de resultados óptimos cada persona tiene un orden de preferencia definido entre ellos y puesto que los resultados son óptimos la preferencia de cada persona debe entrar en conflicto con la preferencia de alguna otra. Como ser racional, cada individuo busca maximizar su propia utilidad y se esfuerza para asegurar ese resultado o esos resultados que prefiere. Suponer que los individuos han de considerar indiferentemente los resultados óptimos, de modo tal que la selección entre ellos sea sencillamente fortuita, es pensar que el proceso del acuerdo carece del carácter dinámico que realmente posee. En este

proceso, algunos resultados pueden surgir de una selección que no depende directa y únicamente de los órdenes de preferencia de los cooperadores individuales. Para que la elección cooperativa sea racional debe reflejar las preferencias de cada persona, pero de qué modo refleja las preferencias es algo que debe depender de la estructura de la interacción que desarrollan los cooperadores y de las consecuencias que ha de tener para cada uno de ellos aquello que es capaz de hacer. Pronto mostraremos cómo varias personas pueden elegir un resultado que ha de alcanzarse mediante una estrategia conjunta, teniendo en cuenta el status de actores que tienen los cooperadores. Al ignorar ese status (bastante razonablemente, si se considera su contexto de aplicación) la elección social se manifiesta irrelevante para la elección cooperativa de una estrategia conjunta.

2.2 Al ser derrotado en su ataque contra el mercado, el utilitarismo contraataca ofreciéndose como una guía para el acuerdo racional.[10] Al principio de la subsección anterior nos preguntábamos si en el caso de la elección cooperativa existe algo análogo a la maximización de la utilidad. El utilitarista responde que realmente hay algo análogo, la *maximización del bienestar* y define el bienestar como la suma de las utilidades individuales. El enfoque del utilitarista difiere del expuesto en 2.1 al defender la posibilidad de que haya una medida de la utilidad individual comparable en el plano interpersonal, lo cual permite sumar las utilidades de las diferentes personas y determinar entonces el bienestar social. La proposición utilitarista identifica la elección cooperativa racional con la maximización de esa suma.

Puesto que la elección cooperativa supone la existencia de un grupo fijo de cooperadores, la proposición de que la suma de las utilidades individuales deba maximizarse es equivalente a la proposición de que debe maximizarse el *promedio* de las utilidades individuales. (El promedio es sencillamente la suma dividida por el número de personas.) Esto le permite al utilitarista sumarse al punto de vista de John Harsanyi, quien demostró que si el bienestar social satisface las mismas exigencias de racionalidad que la utilidad individual (si, en otras palabras, proporciona una medida de intervalo aplicable a los resultados), y si el bienestar que brinda un resultado se relaciona positivamente con las utilidades individuales de ese resultado, el bienestar debe ser pues un promedio mayor de tales utilidades.[11] Además, si el aumento no es arbitrario sino que refleja una medida común de utilidad, el bienestar social debe pues ser sencillamente el promedio de las utilidades individuales. El utilitarista puede entonces afirmar que su proposición es el único método racional de tener en cuenta las preferencias de cada individuo al elegir un resultado que ha de alcanzarse mediante una estrategia conjunta.

Superficialmente, la proposición utilitarista puede parecer muy diferente de la que analizamos en la subsección anterior. En ella comprobamos que todos los resultados óptimos eran indiferentes desde el punto de vista racional. El utilitarista proporciona una medida que los distingue entre sí. Pero ese contraste aparente oculta una afinidad más profunda entre el utilitarismo y la perspectiva de la elección social. Ambas posiciones suponen que la elección de un resultado debe basarse únicamente en las preferencias de las personas afectadas. Ambas posiciones ignoran el proceso de acuerdo existente entre las personas, cada una interesada en maximizar su propia utilidad. La proposición utilitarista, aplicada a la cooperación, considera que los cooperadores son receptores pasivos de los bienes y no seres comprometidos activamente en producir tales bienes y puestos de acuerdo sobre cómo distribuirlos. En la interacción, ninguno de los cooperadores necesita interesarse por el resultado que maximiza la utilidad promedio, salvo la persona que constituye el individuo promedio. Pero aun cuando el individuo promedio sea uno de los cooperadores, es sólo uno entre los muchos que deben participar de la elección cooperativa.

Al analizar el mercado observamos que el bienestar óptimo en general no coincide con el resultado óptimo que se alcanza mediante la competencia perfecta. La redistribución necesaria para hacer que el resultado de la interacción del mercado concuerde con las exigencias utilitaristas puede parecerles irracional a algunas personas que participan de la interacción del mercado. Y podemos suponer que, de manera similar, el resultado óptimo generalmente no coincide con el resultado al que se llega mediante la cooperación. Pues el utilitarista ignora

—como se lo exigen sus principios— la estructura de la interacción. En la perspectiva de la cooperación, el utilitarismo, como una proposición para elegir el resultado, no parece tener nada recomendable.

Por supuesto, no nos oponemos a la demostración de Harsanyi; uno no puede oponerse a las pruebas. Coincidimos en cuanto a afirmar que una medida de intervalo aplicable a los resultados y basada positivamente en las medidas de intervalo individuales de preferencia debe ser utilitarista. Si es posible definir una medida de la utilidad comparable en el plano interpersonal, parece pues que el utilitarismo proporciona el método de elección social exigido desde el punto de vista racional. Pero negamos que la demostración de Harsanyi tenga relevancia en la resolución de nuestro problema. La elección cooperativa racional debe reflejar las preferencias de los cooperadores, pero, aclaremos, sus preferencias como actores. De modo que, sin más rodeos, nos situamos en una perspectiva de la elección cooperativa que toma en cuenta acabadamente tanto la interacción como la preferencia.

2.3 La cooperación surge de la incapacidad que demuestra la interacción del mercado para alcanzar un resultado óptimo a causa de la presencia de las externalidades. De modo que debemos concebir la interacción cooperativa como una mano visible que reemplaza la mano invisible, a fin de alcanzar el mismo ideal que suministra el mercado cuando se desarrolla en condiciones de perfecta competencia. En el intercambio del mercado, los costos y los beneficios se relacionan de un modo que no sólo es óptimo sino que además le proporciona a cada persona un rendimiento de un valor equivalente al de su contribución marginal. La estrategia conjunta seleccionada para lograr una interacción cooperativa debe alcanzar la misma relación de costos y beneficios en condiciones en las que la provisión de un beneficio puede no cubrir sus costos, o en las que la imposición de un costo puede no exigir directamente la provisión de un beneficio compensador. Al aceptar la estrategia cooperativa conjunta como base para sus propias acciones, cada individuo renuncia a las oportunidades de convertirse en un jugador líbero o en un parásito que le brinda la competencia imperfecta, como respuesta a la renuncia que también los demás hacen a tales

oportunidades. Pero, por supuesto, cada individuo se esfuerza por que se elija la estrategia conjunta que le resulte personalmente más favorable, que minimice los costos de su restricción y que maximice los beneficios que ha de recibir de las restricciones a los que están sujetos los demás. Al alcanzar un acuerdo sobre una estrategia conjunta, se ve pues comprometido en un proceso de negociación con sus semejantes. Mediante la negociación, los individuos sientan las bases para una interacción cooperativa que les permite relacionar los costos y los beneficios, aun cuando existan las externalidades, del mismo modo en que lo hace automáticamente el mercado en ausencia de tales externalidades.

La idea de una negociación nos permite incorporar en nuestro punto de vista de la elección cooperativa racional algo que se pierde en las perspectivas de la elección social y el utilitarismo: el compromiso activo de los cooperadores. Además nos permite captar la exigencia de que el acuerdo al que se llegue para encarar una estrategia conjunta sea voluntario. Una negociación racional asegura la participación de cada individuo en la tarea de alcanzar un resultado acordado. Como señalamos al comienzo de este capítulo, no toda cooperación se basa en un acuerdo real. Por lo tanto no debemos suponer que toda estrategia conjunta se elige mediante un procedimiento de negociación del cual participan todos aquellos que basan sus acciones en la estrategia. Pero para que la cooperación sea racional, debemos suponer que la estrategia conjunta fue elegida mediante un procedimiento semejante, de modo tal que cada persona, al reconocerlo, puede aceptar voluntariamente la estrategia. Así cada individuo, en su perspectiva de actor, de participante pleno del proceso cooperativo, puede considerar aceptable la distribución de los beneficios alcanzada mediante la cooperación.

Antes de entrar en el análisis de la negociación racional, deberíamos tomar los recaudos necesario para evitar un malentendido. La interacción cooperativa no es en sí misma una negociación. La interacción cooperativa surge de —y está determinada por— la elección de una estrategia conjunta. De modo que cada individuo elige sus propias acciones según las exigencias marcadas por esa estrategia; al hacerlo ese individuo no

está negociando con sus semejantes. Antes bien podríamos decir que el hecho de elegir esa estrategia conjunta implica una negociación. Por lo tanto la negociación da lugar a la interacción cooperativa, pero es en sí misma no cooperativa. En el análisis que haremos seguidamente esa distinción es de fundamental importancia, pues como veremos, al cooperar, las personas a veces deben restringir su conducta maximizadora de la utilidad, pero en la negociación misma las personas no aceptan tal restricción. Se llega a las restricciones que exige la cooperación mediante la negociación, pero esas exigencias no forman parte del proceso negociador. Para retrotraernos a una distinción que ya debe resultarle familiar al lector a esta altura de la argumentación, digamos que en la negociación la conducta de cada persona debe constituir una respuesta maximizadora de la utilidad a las expectativas que esa persona tiene de la conducta de los demás, mientras que al cooperar la conducta de cada persona debe constituir una respuesta optimizadora a las expectativas que ella se forma de la conducta de las demás sobre la base de la estrategia conjunta elegida por todos.

3.1 La teoría general de la negociación racional es un territorio subdesarrollado. Si bien se han formulado principios de negociación racional con la misma universalidad de aplicación tan libre de contexto como el principio de la maximización de la utilidad esperada, también es cierto que dicha teoría ha sido cuestionada, principalmente por Alvin Roth.[12] John Harsanyi afirma que hay una teoría general y sostiene que él la elaboró basándose en los trabajos anteriores de Frederik Zeuthen y John Nash.[13] Aunque el enfoque Zeuthen-Nash-Harsanyi atrajo el apoyo más amplio de aquellos que aceptan la posibilidad de una teoría general esto no quiere decir que no haya tenido también sus detractores. Animados tanto por el escepticismo de Roth como por el dogmatismo de Harsanyi, nosotros esbozaremos nuestra propia teoría.[14] Los lectores que quieran conocer nuestras razones para preferirla a la teoría de Zeuthen-Nash-Harsanyi podrán encontrarlas en la subsección 3.4.

En cualquier negociación primero es necesario especificar la posición inicial de las partes: la *posición inicial de negociación*, como la llamaremos. Podemos concebir esa posición como un resultado, o como un conjunto de utilidades, una correspondiente a cada negociador. De modo que una situación de negociación consiste en un conjunto de resultados, cualquiera de los cuales los negociadores pueden alcanzar poniéndose de acuerdo sobre la adecuada estrategia conjunta, y un resultado especialmente designado, la posición inicial de negociación. En 1.3 mostramos la representación gráfica del conjunto de resultados: una figura convexa y cerrada que delimita un espacio de utilidad. Por supuesto, la posición inicial de negociación queda representada por un punto interior de ese espacio. La figura 3 ilustra una típica situación de negociación entre dos personas.

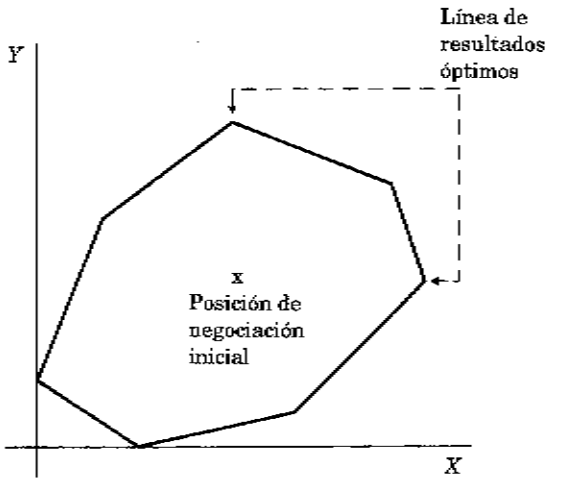


Figura 3

La posición de negociación inicial fija un punto base desde donde parte la negociación. La utilidad que tal posición le proporciona a cada persona representa, en efecto, lo que esa persona aporta a la mesa de negociación y no forma parte de lo que ella busca obtener de la mesa de negociación. Al ponerse de acuerdo en cuanto a adoptar una estrategia conjunta los negociadores están únicamente interesados en la distribución de la utilidad que cada uno puede obtener por encima de lo que obtiene en la posición de negociación inicial. Podríamos decir que los negociadores están interesados en la distribución de las ganancias que puede brindarles la cooperación o en la *plusvalía*

cooperativa. La posición negociadora inicial identifica esa parte de la utilidad de cada individuo que no forma parte de la plusvalía cooperativa.

En 1.3 presentamos la idea de la función de cooperación. Y dijimos que su dominio era un conjunto de espacios-resultados, cada uno de los cuales representa los resultados de una posible interacción cooperativa. Ahora debemos rever ese enfoque a fin de acomodar a él la posición negociadora inicial. El dominio de la función cooperadora es un conjunto de (representaciones de) situaciones de negociación, es decir, un conjunto de espacios-resultados, cada uno de los cuales posee su punto especialmente designado que representa la posición de negociación inicial. En cada situación de negociación, la función de cooperación determina un punto que representa el resultado que ha de alcanzarse mediante la interacción cooperativa. En 1.3 establecimos que ese punto debe caer en el límite superior del espacio-resultado, puesto que debe representar un resultado óptimo. Ahora podemos agregar que ese punto debe representar un resultado que le proporcione a cada persona una utilidad no menor que la utilidad que obtienen en la posición negociadora inicial. En la figura 4 ilustramos el efecto que tiene esa exigencia, al determinar lo que llamaremos la línea de resultados óptimos *admisibles*.

El hecho de que en nuestro enfoque de la función cooperadora incluyamos la posición inicial en la situación negociadora lo distingue formalmente de aquellos enfoques que se basan en la elección social o en el utilitarismo. En esas dos posiciones lo único relevante para elegir un resultado son las preferencias de aquellos individuos implicados (o la medida de tales preferencias). En nuestra perspectiva, las preferencias referentes a una situación particular se seleccionan mediante una consideración especial. Ni la elección social ni el enfoque utilitarista permiten que tales preferencias restrinjan la línea de resultados óptimos disponibles para la selección. Y, como veremos más adelante, la posición negociadora inicial ejerce otros efectos adicionales sobre el proceso de negociación.

Pero, ¿cómo ha de determinarse la posición negociadora inicial? Si suponemos que la cooperación tiene una estructura paralela a la de la interacción del mercado, podríamos proponer

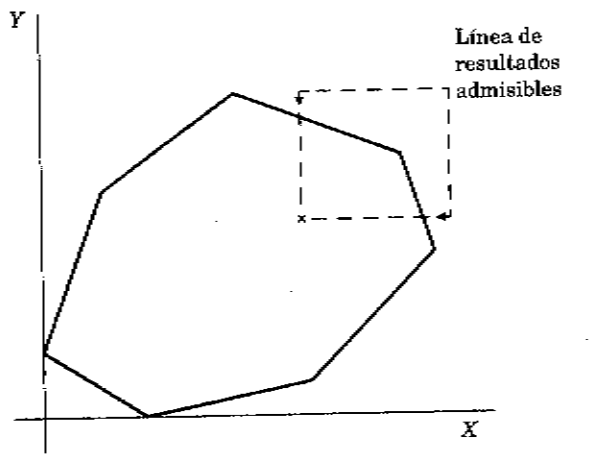


Figura 4

identificar la posición inicial con la distribución inicial de los recursos que han de utilizarse en la cooperación y considerarlos como el conjunto de rendimientos o de utilidades directamente obtenibles de los recursos disponibles iniciales. Pero puede parecer que esta identificación ignora el lugar que ocupa la cooperación en nuestra argumentación. Presentamos la cooperación como una alternativa a la interacción natural y a la interacción de mercado, surgida a causa de la incapacidad de éstas de proporcionar resultados óptimos ante la presencia de las externalidades. Podríamos suponer entonces que el superávit cooperativo es aquello que podría ganarse por encima de lo que resultaría de la interacción no cooperativa. Por consiguiente podemos proponer identificar la posición de negociación inicial con el resultado no cooperativo.

Pero la interacción no cooperativa trae consigo el uso de la fuerza y el fraude, los jugadores líberos y los parásitos. Lo que se intenta con la cooperación es eliminar todos esos males. Por lo tanto para cada negociador, ¿es racional ponerse de acuerdo con sus semejantes si éstos pueden aportar a la mesa de negociación aquello que podrían obtener mediante el juego líbero y el parasitismo o, para expresarlo más correctamente, la utilidad equivalente, y considerar únicamente la distribución de los bienes adicionales que proporciona la cooperación?

Aquí estamos obligados a distinguir dos cuestiones absolutamente diferentes. Por un lado, podemos preguntar qué puede aportar cada individuo a la mesa de negociación, si la cooperación ha de ser racional. Por el otro lado, podemos preguntar qué debe tomar cada individuo de la mesa de negociación, si la cooperación ha de ser racional. Podemos aceptar que cada persona debe obtener de la negociación la expectativa de una utilidad por lo menos igual a la que esperaría de una interacción no cooperativa, si considera racional cooperar. Pero de ello no se sigue que esa persona deba aportar semejante utilidad a la negociación, ni que tal utilidad determine la parte que le corresponde del punto base partiendo del cual se da la negociación.

La determinación de la posición negociadora inicial plantea algunos de los problemas más complejos que debemos examinar. Pero no es necesario examinarlos en este momento; en realidad prefiero diferirlos y tratarlos en el capítulo 7. Tales problemas se refieren a la racionalidad externa de la cooperación. La posición de negociación inicial debe ser fija, de modo tal que para las personas sea racional participar de la cooperación y ponerse de acuerdo en cuanto a llevar adelante una estrategia conjunta y a actuar siguiendo los principios de tal acuerdo. Lo que nos importa aquí es la racionalidad interna de la cooperación, la manera racional de elegir una estrategia conjunta para desarrollar una interacción interdependiente. Por lo tanto, en este capítulo examinaremos la posición negociadora inicial sólo atendiendo a la función que desempeña en la selección de la estrategia cooperativa.

3.2 Puede dividirse el procedimiento de negociación en dos etapas principales. Primero, cada parte presenta su pretensión: propone a la aceptación mutua un resultado o una estrategia conjunta. En general las pretensiones de las partes son incompatibles. De ahí que aparezca la segunda etapa, cada parte —o por lo menos algunas partes— ofrece una concesión renunciando a alguna porción de su pretensión original y propone un resultado alternativo. Salvo en el caso en que haya un desacuerdo insuperable, el proceso de concesión continúa hasta que se logra establecer un conjunto de propuestas mutuamente compatibles. Podemos simplificar esta segunda etapa si suponemos que después de que cada parte presenta sus propuesta inicial,

se alcanza el acuerdo en una sola ronda de concesiones. ¿Qué pretensiones pueden mantener y qué concesiones deben hacer los negociadores racionales?

Cada persona espera que lo que reciba esté en relación con lo que pretende. Cada uno quiere recibir el máximo posible; por lo tanto cada uno pide lo más posible. Pero al decidir hasta dónde puede elevarse el máximo posible, cada individuo se siente restringido porque tiene conciencia de que no puede excluir a los demás de la mesa de negociación ni puede permitir que ellos lo excluyan. De ahí que la pretensión de cada persona esté limitada por el superávit cooperativo general y, más específicamente, por la porción de superávit que tiene la posibilidad de recibir. Pretender más sería proponer que los demás quiten parte de lo que aportaron a la mesa de negociación, parte del rendimiento de la posición negociadora inicial. Y puesto que ninguna persona racional puede esperar que ninguna otra persona racional haga semejante cosa, pretender más que la más amplia porción posible del superávit cooperativo sería vano o, aun peor, pues si alguien quisiera presionar para lograr su pretensión, sólo lograría alejar a los demás de la negociación o que éstos lo expulsen a él. Puesto que cada uno desea beneficiarse compartiendo un superávit cooperativo, nadie tiene interés en provocar el fracaso del proceso negociador, pues su fracaso terminaría con la cooperación y por lo tanto con el superávit. Cada persona al tratar de maximizar su propia utilidad y consciente de que los demás tratan de maximizar la suya, pretende el resultado o la estrategia conjunta que maximice la utilidad que pueda recibir del superávit cooperativo. En otras palabras, cada persona propone , entre los resultados admisibles, aquel que maximice su utilidad. En la figura 5 ilustramos las pretensiones que pueden sostener los individuos racionales en varias situaciones de negociación. Al observarla, parece evidente que en las negociaciones realizadas entre dos personas, los puntos de pretensión quedan determinados fácilmente por el espacio-resultado y por el punto que representa la posición negociadora inicial.

Sin embargo no debemos fiarnos al considerar la negociación entre dos personas, pues eso puede llevarnos a interpretar mal la determinación de las pretensiones. En una situación de

la que participan más de dos personas, cada individuo no ha de pretender siempre todo el superávit cooperativo que pueda recibir, sino aquella parte del superávit que contribuya a producir. La pretensión de cada persona está vinculada con su grado de participación en la interacción cooperativa. Porque si alguien quisiera aplicar presión para obtener aquello que resultaría de la interacción cooperativa de otros, esos otros preferirían excluirlo del acuerdo.

Establecidas las pretensiones de los negociadores, ¿qué concesiones sería racional que hagan? Para poder responder esta pregunta, primero debemos considerar cómo deben medirse las concesiones. La magnitud absoluta de una concesión, atendiendo a la utilidad, es, por supuesto, la diferencia que existe entre la utilidad que un individuo podría esperar del resultado pretendido inicialmente y la utilidad que ese individuo podría esperar del resultado propuesto como concesión. Pero esa magnitud no nos brinda una base para relacionar entre sí las concesiones de los diferentes negociadores, pues la medida de la utilidad individual no permite hacer comparaciones interpersonales.

Sin embargo, podemos establecer una medida de *concesión relativa* que nos permita comparar las concesiones de los diferentes negociadores y que por lo tanto nos suministre una base para determinar qué concesión debe hacer racionalmente cada individuo.

Supongamos que una persona le propone a otra que le haga una concesión y que esta última no esté dispuesta a hacer ninguna concesión (sea la que fuere). Por otro lado, supongamos que una persona fuera a terminar la negociación en su misma posición inicial. En ese caso, esa persona haría en efecto una concesión completa: no recibiría más que lo que aporta a la mesa de negociación, no más de lo que le asignaron las pretensiones de los demás. La *magnitud relativa* de cualquier concesión puede expresarse como la proporción que su magnitud absoluta aporta a la magnitud absoluta de una concesión completa. Si la posición negociadora inicial le proporciona a cierta persona una utilidad u^* y esa persona pretende un resultado que le proporcione una utilidad $u^\#$, si esa persona concede un resultado que le proporcione una utilidad u, la magnitud absoluta de su

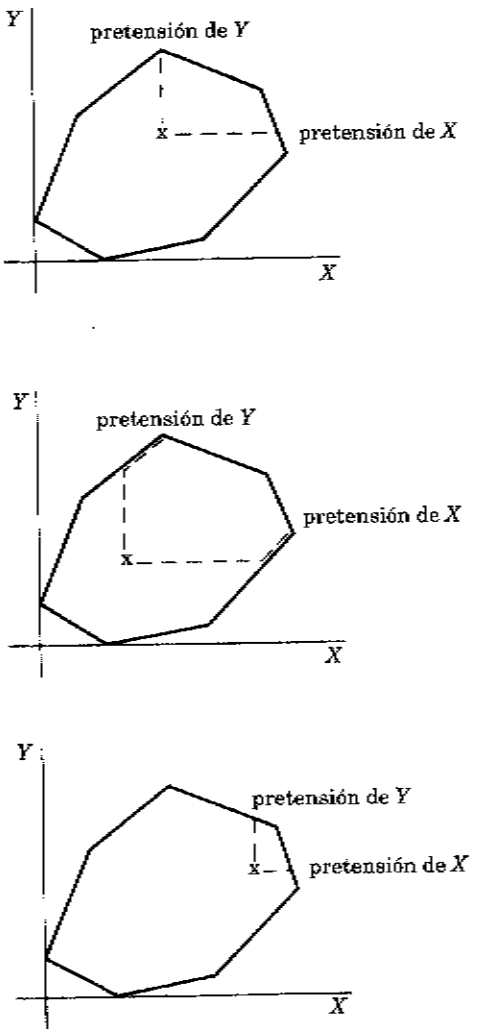


Figura 5

concesión es $(u^\# - u)$ de la concesión completa $(u^\# - u^*)$ y por lo tanto la magnitud relativa de su concesión es $[(u^\# - u)/(u^\# - u^*)]$.

La concesión relativa es una proporción de dos diferencias o dos intervalos de utilidad. Ahora una medida de intervalo, tal como la utilidad, fija esas proporciones independientemente de sus rasgos arbitrarios: el punto cero y la unidad. Consideremos

tres temperaturas como se las denomina según las escalas Celsius y Fahrenheit: sean 10° C = 50°F, 20°C = 68°F y 25°C = 77°F. Medida en la escala Celsius la diferencia o el intervalo entre la primera y la segunda es de 10°C, y entre la segunda y la tercera, es de 5°C. Medida en la escala Fahrenheit, los intervalos son de 18 y de 9°F. Pero la proporción entre los dos intervalos 10/5 y 18/9 es 2, independientemente de la escala elegida. El ejemplo de la temperatura ilustra algo que es válido para todas las medidas de intervalos y por lo tanto para la utilidad. La concesión relativa es independiente de la elección de una escala de utilidad. De modo que las concesiones relativas de cada persona son fijas, sea como fuere que elijamos medir sus utilidades.

Además hemos presentado la concesión relativa de modo tal que para cada persona la magnitud relativa de no hacer ninguna concesión es siempre 0, y la magnitud relativa de una concesión completa o total es siempre 1. Por lo tanto tenemos una medida de concesión relativa que, sin introducir ninguna comparación de utilidad interpersonal (porque no debemos suponer que a concesiones relativas iguales correponden costos de utilidad iguales) nos permite sin embargo comparar las concesiones que piden las diferentes personas que participan de una situación de negociación.

Ahora representamos cada resultado admisible no desde el punto de vista de las utilidades que le brinda a cada persona, sino desde el punto de vista de las concesiones que exigiría que cada uno haga, si se lo eligiera para determinar la estrategia conjunta que se aplicaría en la interacción cooperativa. Entonces puede emplearse el principio de Zeuthen para establecer una regla que decida qué concesiones pueden hacerse racionalmente y por lo tanto qué resultado debe ser aceptado. El lector recordará de 3.2.2 que ese principio establece que la persona cuyo promedio entre el costo de la concesión y el costo del desacuerdo insuperable sea menor debe conceder al otro. O para utilizar los términos del presente análisis, el principio establece que la persona con una menor concesión relativa debe conceder. Si extendemos los alcances de esta regla a la negociación entre varias personas, podemos decir que el principio debería establecer que dados varios resultados posibles, cada uno de los cuales,

para ser seleccionado exige que alguna persona o todas las personas hagan ciertas concesiones, un resultado sólo ha de elegirse si la concesión relativa mayor o *máxima* que exige es lo menor posible, o un *mínimo*, es decir, que no sea mayor que la concesión relativa máxima exigida por ninguno de los otros resultados. Llamamos a este principio, el principio de mínimo-máximo o de *concesión relativa minimáxima*.

Volvamos a un ejemplo familiar: el de Jane y Brian. Supongamos que identificamos la posición negociadora inicial con el único resultado en equilibrio que ellos pueden esperar que se dé de una interacción no cooperativa; ese resultado le proporciona a Jane una utilidad de 1/2 y a Brian una utilidad de 3/8. Esta situación queda ilustrada en la figura 6. Los resultados admisibles son aquellos resultados óptimos que le proporcionan a Jane una utilidad de por lo menos 1/2 y a Brian una utilidad de por lo menos 3/8; los puntos que representan tales resultados caen en la línea que une (0, 1) y (1,0), entre los puntos (1/2, 1/2) y (5/8, 3/8) inclusive. Es evidente que Jane pretende el segundo y Brian el primero. El resultado que le brinda a Jane una utilidad de 9/16 y a Brian una utilidad de 7/16 exige de Jane una concesión relativa de [(5/8 -9/16)/ (5/8 - 1/2)] = 1/2. Ningún resultado admisible más favorable a Jane ha de exigir a Brian una concesión relativa mayor, y ningún resultado más favorable a Brian ha de exigir a Jane una concesión relativa mayor. De ahí que por el principio de concesión relativa minimáxima, Jane y Brian deberían ponerse de acuerdo respecto de ese resultado que le proporciona a Jane 9/16 y a Brian 7/16; si ambos cooperan, Jane debería ir a la fiesta y Brian debería seguir una estrategia mixta que le asigne una probabilidad de 7/16 a asistir a la fiesta y de 9/16 a permanecer en su casa.

Un segundo ejemplo ilustra cómo ciertas objeciones aparentemente plausibles al principio de concesión relativa minimáxima, tomado como base para la cooperación, pueden rebatirse. Adelaide y Ernest tienen una oportunidad de cooperar de un modo mutuamente beneficioso si primero pueden ponerse de acuerdo sobre cómo dividir las ganancias. Adelaide recibiría un beneficio neto de $500 de la empresa conjunta, siempre que reciba toda la ganancia después de que se cubran completamente los costos de Ernest; sin embargo Ernest recibi-

ría sólo $50 si percibe todas las ganancias después de que se cubran todos los costos de Adelaide. Puesto que los bienes actuales de ambos no están en disputa, y puesto que ninguno de los dos tiene una fuente alternativa de ganancia, podemos establecer que lo que cada uno aporta a la situación inicial de negociación es $0. Suponemos que los resultados óptimos admisibles caigan en la curva que se muestra en la figura 7.

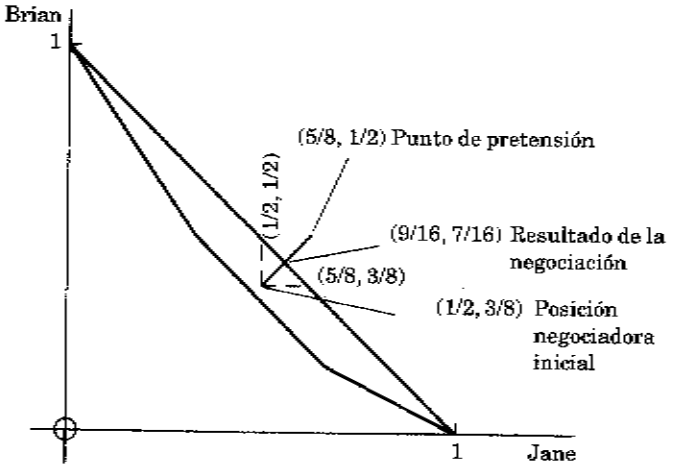


Figura 6

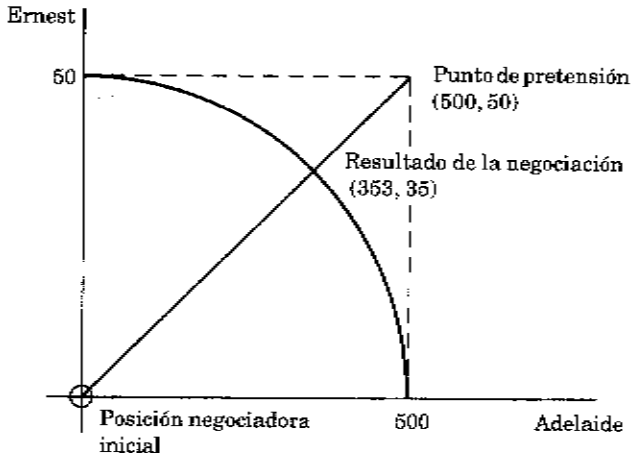


Figura 7

Cada uno pretende tanto como le es posible siempre que le proporcione al otro un aliciente marginal para cooperar, de modo tal que la pretensión de Adelaide se aproxime a $500 y la de Ernest a $50. En la figura 7 vemos que el acuerdo es posible si cada uno hace una concesión relativa de aproximadamente 0,3, que proporciona un resultado que le da a Adelaide una ganancia de $353 y a Ernest una de $35. También vemos que ponerse de acuerdo sobre cualquier otro resultado exigiría que uno de los dos haga una concesión relativa mayor. De ahí que Adelaide deba conceder $147 de su ganancia posible y aceptar $353; Ernest debe conceder $15 de su ganancia posible y aceptar $35.

Supongamos que Ernest se queje de que Adelaide reciba mucho más que él: que compare los $353 con sus meros $35. Adelaide respondería que Ernest cede mucho menos que ella, compararía los meros $15 con los $147 que ella deja de recibir. Si Ernest quisiera defender la posición de que debería aumentarse su ganancia porque ésta es mucho menor que la de Adelaide, Adelaide podría sostener que lo que debería aumentarse es la concesión que hace Ernest porque es mucho menor que la que hace ella. Si Adelaide quisiera sostener la posición de que Ernest se preocupa relativamente poco por lo que cede, Ernest podría replicarle que se preocupa igualmente poco por lograr un acuerdo. El beneficio potencialmente menor que le corresponde a Ernest reduce la presión que siente para rechazar una concesión, pero igualmente reduce su interés por lograr un acuerdo y por lo tanto su interés por hacer una concesión. El mayor valor de lo que está en juego para Adelaide aumenta la presión que ella siente por alcanzar el acuerdo, pero igualmente aumenta la presión para oponerse a una concesión comparable en magnitud relativa.

La racionalidad fundamental del principio de concesión relativa minimáxima no exige siquiera las burdas comparaciones interpersonales del beneficio que hemos utilizado en este ejemplo. Antes bien, la racionalidad se refiere a una comparación interpersonal de la proporción de la ganancia potencial que cada persona debe conceder. Sin embargo, si tuviéramos que imaginar una medida de la utilidad que permita hacer comparaciones interpersonales y si nos sintiéramos tentados a aplicar

algún principio de ganancia igual, deberíamos recordar que ninguna tentación semejante tiene como contrapartida un principio de pérdida igual, en relación con la pretensión de cada uno. El hecho de que la concesión relativa minimáxima es la única aceptable se hace evidente cuando equilibramos la ganancia en relación con la no cooperación y la pérdida en relación con la pretensión o la ganancia potencial.

Los dos ejemplos que hemos analizado pueden indicar que el principio de concesión relativa minimáxima es en realidad un principio de concesión relativa igual mínima, es decir, que exige de los negociadores las menores concesiones iguales medidas relativamente. Ya defendí antes esta posición, pero equivocadamente. En muchos casos, la concesión relativa minimáxima puede ser equivalente a concesiones relativas iguales, pero Alvin Roth demostró que no siempre es así.[15]

Lo que podemos mostrar es que, si hay un resultado óptimo que exige que todas las personas hagan concesiones relativas iguales, el principio de concesión relativa minimáxima elige ese resultado. Porque si un resultado es óptimo, todo otro resultado de la situación debe proporcionarle a alguna persona una utilidad menor. Pero, para aceptar un resultado que brinde una utilidad menor, un individuo debe hacer una concesión mayor, mayor en magnitud absoluta, puesto que la diferencia entre la utilidad de la pretensión de ese individuo y la utilidad de la concesión que hace aumenta y por lo tanto, es más grande en magnitud relativa, puesto que la proporción que aporta esa diferencia a la concesión completa también debe aumentar. Por consiguiente, si hay un resultado óptimo que exige que todas las personas hagan concesiones relativas iguales, cualquier otro resultado exige una concesión mayor que el minimáximo.

Se puede representar gráficamente de manera sencilla el resultado que exige una concesión relativa minimáxima, dado que hay un resultado óptimo que exige a cada uno de los individuos concesiones relativas iguales. Supongamos que el punto de pretensión sea el punto situado dentro del espacio de utilidad que representa, para cada persona, una utilidad igual a la que podría obtener de su pretensión. Entonces, todos los puntos situados en la línea que une el punto que representa la posición negociadora inicial con el punto de pretensión, repre-

sentan resultados que exigen la misma concesión relativa. Si esa línea intercepta el límite superior derecho del espacio-resultado, el punto de intersección representa pues el resultado que exige las mínimas concesiones relativas iguales, que es el resultado seleccionado por el principio de concesión relativa minimáxima. Mostramos esto en las figuras 6 y 7.

Un tercer ejemplo vincula el principio de concesión relativa minimáxima con la suposición evidente de que en una sociedad, cada individuo debería beneficiarse en proporción con lo que invierte. Abel y Mabel se ponen de acuerdo para unir sus recursos y lo consideran ventajoso porque por encima del monto de sus contribuciones hay un aumento lineal de retribución que está en relación con la mayor inversión. Los recursos de ambos sumados tienen un valor de 1 y ellos esperan una retribución valuada en k (como en el ejemplo anterior, suponemos que las utilidades pueden tratarse como valores monetarios). Si Abel contribuye con r, la contribución de Mabel a los recursos combinados es $(1-r)$ y si Abel, invirtiendo independientemente, pudiera esperar una retribución valuada en c, basándonos en que hay un aumento lineal de la retribución proporcional a la mayor inversión, decimos que Mabel debería esperar una retribución valuada en $[(k - 2kr + cr)/(1 - r)]$ si invirtiera independientemente. Abel recibiría cr si invirtiera independientemente y Mabel recibiría $(k - 2kr + cr)$. Decimos que tales retribuciones son las que obtendrían de la posición negociadora inicial. La pretensión de Abel es pues k menos la retribución inicial de Mabel, o $(2kr - cr)$ y de manera similar, la pretensión de Mabel es $(k - cr)$.

En esta situación, el principio de concesión relativa minimáxima exige de Abel y de Mabel concesiones relativas iguales. Si decimos que la retribución que recibe Abel de la cooperación es x, la concesión absoluta de su pretensión sería pues $(2kr - cr)$ menos x; es decir, la concesión plena de su pretensión menos su retribución inicial; de modo que su concesión relativa es $[(2kr - cr - x)/(2kr - 2cr)]$. De manera similar, la concesión relativa de Mabel es $[(x - cr)/(2kr - 2cr)]$, puesto que la retribución que reciba de la cooperación debe ser $(k - x)$. Puesto que las concesiones relativas de ambos son iguales, podemos igualarlas y resolver x. Comprobamos que $x = kr$. La retribución kr que obtiene Abel es igual a la retribución total k multiplicada

por la proporción que ocupa su inversión r en la inversión total 1. Y la retribución k de Mabel $(1-r)$ es igual a la retribución total multiplicada por la proporción que ocupa la inversión de Mabel en la inversión total. Por supuesto, estos resultados son los que deberíamos esperar si la concesión relativa minimáxima suministra, como lo afirmamos, la base racional para la cooperación.

3.3 Ahora daremos una explicación formal de la teoría de la negociación racional esbozada ya en las subsecciones anteriores. Comenzaremos dando algunas definiciones:

(1) La *posición de negociación inicial* es un resultado, cuyas utilidades constituyen las dotes negociadoras de los futuros cooperadores. (Por lo tanto, cada negociador tiene derecho a la utilidad que le estaba garantizada en la posición negociadora inicial y que no está en juego durante el proceso de negociación.)
(2) El *superávit cooperativo* es un conjunto de diferencias de utilidad, una correspondiente a cada cooperador, ninguna de valor de negativo y todas iguales a la diferencia que existe entre la utilidad que proporciona la cooperación y la utilidad que garantiza la posición negociadora inicial. (La distribución del superávit cooperativo, es decir, la selección de un conjunto particular de diferencias de utilidad que constituye tal superávit, está en juego en el proceso de negociación.)
(3) Una *situación de negociación* es un conjunto de resultados, representados en un espacio-utilidad por una figura convexa y cerrada, el *espacio-resultado*, y una posición negociadora inicial, representada por un punto del espacio-resultado.
(4) Una *pretensión* es aquello que demanda un futuro cooperador de un particular superávit cooperativo en el momento de iniciar la negociación.
(5) El *punto de pretensión* es el punto situado en el espacio de utilidad que representa el resultado (posible hipotético) que podría proporcionarle a cada persona una utilidad igual a la de su pretensión; ese resultado es *factible* si y sólo si es un punto situado en el espacio-resultado.
(6) Una *concesión* es un ofrecimiento que hace un futuro cooperador de aceptar una utilidad particular menor que la que pretendía.
(7) Un *punto de concesión* es un punto situado en el espacio de utilidad que representa el resultado (posible hipotético) que surgiría de un conjunto de concesiones, una hecha por cada futuro cooperador (pero que posiblemente incluya concesiones nulas), ese resultado es *factible* si y sólo si es un punto situado dentro del espacio-resultado.

(8) Una persona esta dispuesta a *tomar en consideración* un punto de concesión si y sólo si está dispuesta a hacer la concesión que se le pide siempre que los demás estén dispuestos a hacer las concesiones que se les pide. La persona está pues dispuesta a tomar en consideración la concesión en relación con el punto de concesión.
(9) La *magnitud absoluta* de una concesión es la diferencia que hay entre la utilidad de la pretensión de la persona y la utilidad de la concesión.
(10) La *magnitud relativa* de una concesión es la proporción que ocupa su proporción absoluta en la diferencia que existe entre la utilidad de la pretensión de la persona y la utilidad que le garantiza su posición negociadora inicial. Alternativamente, la magnitud relativa de una concesión es la proporción que ocupa su magnitud absoluta en la diferencia de utilidad que le brinda a la persona el superávit cooperativo que pretende.
(11) Una *concesión máxima* para cualquier punto de concesión es la concesión (o una de las concesiones) que tiene la mayor magnitud relativa exigida para alcanzar el resultado representado por el punto.
(12) Una *concesión minimáxima* en cualquier situación de negociación es la concesión máxima (o una de las concesiones máximas) que tiene la magnitud relativa mínima exigida para alcanzar un resultado representado por un punto de concesión factible. (Por consiguiente en cualquier situación de negociación, cada resultado representado por un punto de concesión factible debe exigir una concesión que tenga una magnitud relativa por lo menos igual a aquella de la concesión minimáxima.)

Ahora formularemos las condiciones de la negociación racional:

(*i*) La *pretensión racional*. Cada persona debe pretender el superávit que le proporcione la máxima utilidad, salvo que ninguna persona pueda pretender un superávit cooperativo si no fuera un participante de la interacción necesaria para alcanzarlo.
(*ii*) El *punto de concesión*. Dadas las pretensiones que satisfagan la condición (*i*), cada persona debe suponer que existe un punto de concesión factible que toda persona racional esté dispuesta a tomar en consideración.
(*iii*) La *disposición a conceder*. Cada persona debe estar dispuesta a tomar en consideración una concesión en relación con un punto de concesión si la magnitud relativa de tal concesión no es mayor que la de la concesión mayor que el individuo supone

que una persona racional está dispuesta a tomar en consideración (en relación con un punto de concesión factible).
(*iv*) Los *límites de la concesión*. Ninguna persona está dispuesta a tomar en consideración una concesión que esté en relación con un punto de concesión, si no se lo exigen las condiciones (*ii*) y (*iii*).

La racionalidad de estas condiciones se refiere al beneficio que cada persona busca obtener del superávit cooperativo. Cada individuo puede aumentar su utilidad mediante la cooperación, de ahí que, como maximizador de la utilidad, cada individuo debe considerar racional el hecho de cooperar. Y cada uno reconoce que todos los demás también deben considerarlo racional.

La condición (*i*) es una aplicación directa de la maximización de la utilidad al contexto de la negociación. Cada individuo trata de maximizar el rendimiento que le proporciona la cooperación, de modo tal que cada uno pretende lo máximo posible, pero nadie quiere excluir a los demás cooperadores ni que ellos lo excluyan, por ello cada individuo limita sus pretensiones para impedir esas exclusiones.

La condición (*ii*) procede del hecho de que, para que se dé la cooperación, todos deben ponerse de acuerdo respecto de un resultado (o de una estrategia conjunta) representada por un punto factible del espacio-resultado. Cualquier persona que suponga que no hay un punto de concesión factible sobre el cual las personas racionales puedan ponerse de acuerdo estará negando que haya alguna forma de que esas personas racionales cooperen. Pero todos reconocen que todos consideran racional el hecho de cooperar.

La condición (*iii*) expresa la igual racionalidad de los negociadores. Puesto que cada persona, como maximizadora de la utilidad, trata de minimizar su concesión, nadie puede pues esperar que ninguna persona racional esté dispuesta a hacer una concesión si no está dispuesto a su vez a hacer una concesión similar.

Finalmente, la condición (*iv*) constituye también una aplicación directa de la maximización de la utilidad, dado que las demás condiciones son suficientes para exigir concesiones que lleven a obtener un resultado representado por un punto factible. Ninguna persona racional puede estar dispuesta a hacer concesiones innecesarias o concesiones innecesariamente amplias. Ahora demostramos que las condiciones (*i*) y (*ii*) bastan para ponerse de acuerdo respecto de un punto factible.

Al expresar la definición (12) hicimos notar entre paréntesis que cada resultado representado por un punto de concesión factible debe exigir una concesión con una magnitud relativa por lo menos igual a la concesión minimáxima. Puesto que por la condición (*ii*) cada persona debe suponer que existe un punto de concesión factible que toda persona racional está dispuesta a tomar en consideración, cada individuo debe pues suponer que toda persona racional está dispuesta a tomar en consideración un punto de concesión que exija que alguien haga una concesión cuya magnitud relativa sea por lo menos igual a la concesión minimáxima, en relación con ese punto de concesión. De la condición (*iii*) se sigue que cada persona debe estar dispuesta a tomar en consideración una concesión que tenga una magnitud relativa por lo menos igual a la concesión minimáxima, en relación con un punto de concesión factible.

Pero, en cada situación hay un punto de concesión factible que exige que ninguna concesión tenga una magnitud relativa mayor que la concesión minimáxima. Por lo tanto, por las condiciones (*ii*) y (*iii*) toda persona debe estar dispuesta a tomar en consideración ese punto. Pero entonces la condición (*ii*) no puede exigirle a ninguna persona que suponga que existe un punto de concesión factible que toda persona racional esté dispuesta a tomar en consideración y que exige una concesión de mayor magnitud relativa que la concesión minimáxima. De modo que la condición (*iii*) no puede, ni por sí misma ni basándose en la condición (*ii*), exigirle a ninguna persona que esté dispuesta a tomar en consideración una concesión cuya magnitud sea mayor que la concesión minimáxima, en relación con cualquier punto de concesión factible. De ahí que por la condición (*iv*) ninguna persona esté dispuesta a tomar en consideración una concesión de magnitud mayor que la concesión minimáxima. Así, cada persona debe estar dispuesta a tomar en consideraciones aquellos y sólo aquellos puntos de concesión factibles que exigen concesiones cuyas magnitudes relativas sean tan grandes (pero no mayores) que la concesión minimáxima.

Digamos que las pretensiones que satisfacen la condición (*i*) son *máximas*. Partiendo de las condiciones (*i*) y (*iv*) establecimos el *Principio de concesión relativa minimáxima*: en cualquier interacción cooperativa, la estrategia conjunta racional queda determinada por una negociación llevada a cabo entre los cooperadores, negociación en la cual cada individuo presenta su pretensión máxima y luego ofrece una concesión cuya magnitud relativa no puede ser mayor que la concesión minimáxima.

El principio de concesión relativa minimáxima desempeña una función triple en nuestra argumentación. Primero, expresa el principio de maximización de la utilidad esperada en el contexto de la negociación. Los negociadores racionales, cada uno de los cuales trata de maximizar su propia utilidad, determinan sus pretensiones y sus concesiones apelando a este principio. En segundo lugar, el principio determina el contenido formal de una negociación racional; los negociadores racionales se ponen de acuerdo sobre esa estrategia conjunta que le proporciona a cada persona una utilidad esperada no menor que la utilidad que esa persona podría esperar de su pretensión máxima y su concesión minimáxima. De modo que el principio domina tanto el proceso como el objeto de la elección racional en situaciones de negociación.

En tercer lugar, el principio de concesión relativa minimáxima es el principio de la conducta racional en la interacción cooperativa, la interacción basada en la estrategia conjunta aceptada en la negociación. Cada individuo obra no para maximizar su propia utilidad sino para alcanzar el resultado que es el objeto de la negociación y que le proporciona a cada persona una utilidad esperada no menor que la que podría esperar de su pretensión máxima y su concesión minimáxima. Aún debemos demostrar que tal acción es racional, que racionalmente cada persona debería obrar de acuerdo con la estrategia conjunta que aceptó racionalmente. Esa será nuestra tarea del próximo capítulo. Esta tercera función del principio es la que establece el carácter distintivamente moral del principio y de la cooperación. Porque, aplicado a la interacción cooperativa, el principio de concesión relativa minimáxima constituye una restricción a la búsqueda directa de la utilidad individual. Por consiguiente, si podemos mostrar que este principio constituye

una base racional e imparcial para la interacción cooperativa, tendremos el fundamento necesario para considerarlo un principio moral.

Al tener que afrontar los costos de la interacción natural o de la interacción del mercado empeñada por las externalidades, las personas racionales aceptan una forma de interacción cooperativa diferente. Aceptar obrar, no sobre la base de la maximización de la utilidad individual, sino antes bien sobre la base de la optimización, en la cual el principio de concesión relativa minimáxima determina el resultado óptimo particular. Por supuesto, al alcanzar ese acuerdo, cada individuo busca maximizar su propia utilidad. El mismo principio de concesión relativa minimáxima les sirve a las personas racionales tanto para alcanzar el acuerdo como para ajustarse a él —tanto como un principio de elección en la negociación, como un principio de elección en la cooperación—. En la negociación, la concesión relativa minimáxima expresa directamente las demandas de maximización de la utilidad, como ya lo sostuvimos en la presente sección. En la cooperación, la concesión relativa minimáxima restringe y por lo tanto pasa por alto las demandas de maximización de la utilidad. Este será el tema del próximo capítulo.

3.4 Completaremos el tratamiento de este tema con un breve análisis de la principal posición alternativa a nuestro enfoque de la negociación racional, la teoría Zeuthen-Nash-Harsanyi.[16] Los lectores cuyo interés no sea tan grande como para querer lidiar con los teóricos de la negociación pueden omitir esta subsección pues no contiene nada de importancia positiva para el desarrollo de nuestra argumentación.

Consideremos dos personas, Ann y Adam, quienes se encuentran en una situación de negociación. Supongamos que la posición negociadora inicial le proporciona a Ann una utilidad u^* y a Adam una utilidad v^*. Ann pretende un resultado que le proporcione una utilidad u_1 y a Adam una utilidad v_1, Adam pretende un resultado que le proporcione una utilidad v_2 y que le dé a Ann una utilidad u_2. Suponiendo que las pretensiones de ambos sean incompatibles, es decir que (u_1, v_2), el punto que les proporcionaría las utilidades pretendidas, no sea factible, por lo menos uno de los dos debe ofrecer hacer una concesión. Si Ann

aceptara la pretensión de Adam, recibiría $(u_1 - u_2)$; si mantuvieran un desacuerdo insuperable en la posición negociadora inicial, Ann recibiría $(u_1 - u^*)$. Si Adam aceptara la pretensión de Ann recibiría $(v_2 - v_1)$; si ambos mantuvieran un desacuerdo insuperable en la posición negociadora inicial, Adam recibiría $(v_2 - v^*)$. La razón $[(u_1 - u_2)/(u_1 - u^*)]$ mide la pérdida que sufre Ann si acepta la pretensión de Adam como una proporción de lo que pierde si hay un desacuerdo insuperable; de manera similar la razón $[(v_2 - v_1)/(v_2 - v^*)]$ mide la pérdida que sufre Adam si acepta la pretensión de Ann como una proporción de lo que pierde si hay un desacuerdo insuperable. Podemos comparar los resultados de esas razones y suponer, aplicando el principio de Zeuthen, que la persona a la que le corresponde una razón menor arriesga más al no hacer una concesión o, alternativamente, pierde proporcionalmente menos haciendo una concesión.

Pero esa persona no necesita hacer una concesión completa y aceptar la pretensión del otro, si, al hacer una concesión menor, puede aumentar suficientemente su razón o proporción, de modo tal que ésta sea mayor que la del otro negociador. Supongamos que Ann deba conceder; $[(u_1 - u_2)/(u_1 - u^*)]$ es menor que $[(v_2 - v_1)/(v_2 - v^*)]$. Pero supongamos que Ann pueda proponer un resultado que le proporcione una utilidad, u_3 y le dé a Adam una utilidad v_3, en el cual u_3, aunque menor que la pretensión original u_1, es más de lo que Ann recibiría aceptando la pretensión de Adam que le asigna una utilidad u_2, y en el que también $[(u_3 - u_2)/(u_3 - u^*)]$ es mayor que $[(v_2 - v_3)/(v_2 - v^*)]$. Ahora el riesgo que corre Adam es mayor que el que corre Ann, de modo que él debe ofrecer una concesión.

Suponemos que este proceso de concesión continúa hasta que las proposiciones de ambos negociadores coinciden. Puede mostrarse que si en la negociación se desarrolla este proceso, el resultado seleccionado debe proporcionarle a Ann una utilidad u' y a Adam una utilidad v', tales que para cualquier punto factible (u, v), el producto $(u' - u^*)(v' - v^*)$ es por lo menos tan grande como el producto $(u - u^*)(v - v^*)$. De modo que, según la teoría de Zeuthen-Nash-Harsanyi, el acuerdo racional maximiza el producto de las diferencias de utilidad individuales en el superávit cooperativo.

En alguna situación este procedimiento ha de proporcionar el mismo resultado, como el que propusimos nosotros. Pero en la figura 8 mostramos una situación negociadora en la que el procedimiento difiere. Los resultados posibles están representados por los puntos (0, 0), (1/2, 1) y (1, 0); (0, 0) es (suponemos) la posición negociadora inicial. Los resultados óptimos están representados por los puntos situados en la línea que une (1/2, 1) con (1, 0). Por el procedimiento que acabamos de describir, el resultado de la negociación debe estar representado por ese punto situado en la línea óptima maximizadora del valor de $(u - 0)(v - 0)$; es decir, el punto (1/2, 1).

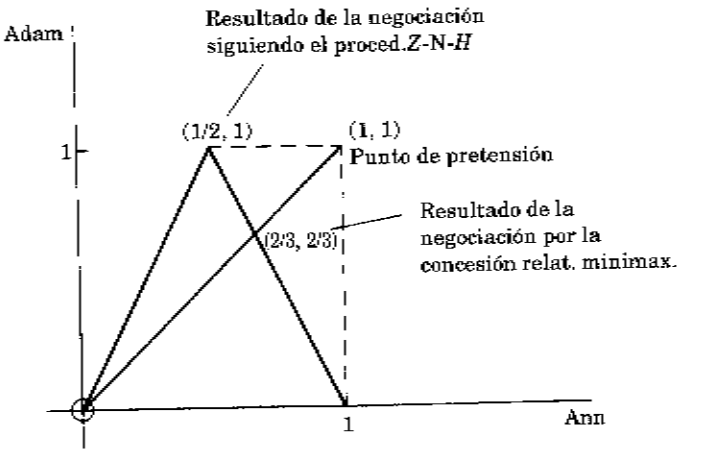


Figura 8

Siguiendo nuestro procedimiento, está claro que tanto Ann como Adam deben pretender una utilidad de 1; el punto (2/3, 2/3) sobre la línea óptima exige que cada negociador haga una concesión con una magnitud relativa $[(1 - 2/3)/(1 - 0)]$, o 1/3, y ningún otro punto, puesto que le proporciona a Ann o a Adam una utilidad menor que 2/3, le exige a uno de ellos que haga una concesión con una magnitud relativa mayor. Por lo tanto, por el principio de concesión relativa minimáxima, el resultado de la negociación queda representado por el punto (2/3, 2/3).

No sostenemos que la intuición dicte la última palabra al decidir las cuestiones de la teoría de la elección racional. Pero si

la intuición tiene algo que decir, esas palabras no están aquí para sostener el enfoque de Zeuthen-Nash-Harsanyi que le permite mantener a Adam su pretensión y deja que Ann haga todas las concesiones.

Pero, permítasenos hacer algunas objeciones más profundas. Según el procedimiento de Zeuthen-Nash-Harsanyi, el resultado queda determinado por la maximización del producto de las diferencias de utilidad individuales que hay en el superávit cooperativo. La única restricción que debe aplicarse a las pretensiones de los negociadores para asegurar ese resultado, es que éstos especifiquen los resultados que son admisibles en el sentido de 3.1, y que le proporcionen al individuo no menos de lo que recibiría del resultado de la negociación. Mientras que siguiendo nuestro procedimiento, cada persona tiene un interés al presentar su pretensión máxima, en esta alternativa ninguno tiene semejante interés, puesto que nadie puede esperar más pretendiendo más. Es decir que no hay una racionalidad que determine las dimensiones de la pretensión de cada individuo.

En cada etapa del proceso de negociación se hace una concesión mínima, en concordancia con la exigencia de la racionalidad maximizadora. Pero, ¿una concesión de qué? De una pretensión que carece de significación. No es posible determinar, al final del proceso negociador, si una parte concedió más que la otra, puesto que las concesiones se miden únicamente en cada etapa del proceso y son relativas a esa etapa. Y aun cuando existiera una sola etapa, podemos decir que una parte concedió más sólo en relación con las pretensiones iniciales que, como ya vimos, se establecieron de un modo completamente arbitrario. Seguramente debemos preguntarnos pues si la idea de concesión conserva una verdadera significación en este enfoque de la negociación. Al considerar que la magnitud de las pretensiones es arbitraria, hacemos que también la magnitud de las concesiones llegue a ser arbitraria.

Por supuesto, podríamos suponer que los negociadores han de comenzar estableciendo pretensiones reales, pretensiones que satisfagan nuestra condición (*i*). Aunque no tenga razón para hacerlo, si lo hacen, eso nos permitiría atribuirle significación a la medida de las concesiones de cada persona, tomando como punto de partida esas pretensiones. Pero entonces deberíamos afrontar la evidente y desagradable consecuencia de que, así medidas, las concesiones no tienen por qué ser mínimas. Puesto que el resultado del procedimiento de negociación de Zeuthen-Nash-Harsanyi no necesita ser el resultado al que se llega según nuestro procedimiento, podemos estar seguros de que aquél no siempre ha de satisfacer el principio de concesión relativa minimáxima. Entonces podemos preguntarnos por qué los negociadores racionales habrían de emplear un procedimiento que le exigiría a uno de ellos concesiones mayores que las que ninguno tiene necesidad de hacer.

La médula de esta teoría alternativa de la negociación se funda en uno de los postulado de Nash, generalmente llamado "independencia de las alternativas irrelevantes". El postulado, expresado atendiendo a la representación de la situación negociadora en el espacio de utilidad, establece que el hecho de restringir el espacio-resultado sólo afecta el resultado de la negociación si hace que el punto que representa el resultado original no sea factible, fuera del espacio-resultado restringido, o si afecta la posición negociadora inicial. Alternativamente, expandir el espacio-resultado sólo puede afectar el resultado de la negociación si el nuevo resultado cae fuera del espacio-resultado original o si, también en este caso, afecta la posición negociadora inicial. Este postulado expresa de manera explícita la irrelevancia de las pretensiones de los negociadores, puesto que restringir o ampliar el espacio-resultado claramente puede afectar las pretensiones posibles sin afectar el resultado de la negociación. Por consiguiente, el postulado de Nash niega lo que nosotros afirmamos, es decir que el resultado de la negociación racional depende de las pretensiones establecidas por los negociadores.

Harsanyi utiliza un conjunto diferente de postulados, entre los cuales incluye la exigencia plausible de que el resultado de la negociación no debe depender (o no debe esperarse que dependa) de factores cuya relevancia no pueda establecerse sobre la base de otros postulados.[17] Seguramente en sí misma esta exigencia es aceptable. Pero, por la forma en que describe el proceso negociador, Harsanyi la utiliza para llegar a la conclusión de que sólo aquellos factores que forman parte de la elección entre la pretensión de un individuo y la pretensión de

otro —entre hacer y no hacer una concesión— son relevantes. Si uno supone que las personas plenamente racionales actúan mediante una serie de pretensiones (o de ofrecimientos) y concesiones que conducen a establecer nuevas pretensiones y consideran que cada paso implica una evaluación comparativa de los riesgos inmediatos, no es sorprendente que llegue a la misma conclusión de Harsanyi. Pero si uno supone que las personas racionales obran estableciendo pretensiones máximas y haciendo luego un sólo conjunto de concesiones, es natural que llegue a nuestra conclusión. Porque entre las dos descripciones del proceso varía aquello que consideramos relevante para la negociación.

Uno podría suponer que el carácter seriado de la negociación (como lo representa Harsanyi) es un enfoque más acertado de la verdadera naturaleza de la negociación que el que damos nosotros. Pero, como hemos visto, si se supone que ese carácter seriado de pretensión-concesión-pretensión es central, se comprueba que desde el punto de vista racional el proceso mismo es algo fingido. Si no importa lo que se pretende, lo que uno concede se basa en algo que no tiene importancia. El resultado al que se llega según el procedimiento de Zeuthen-Nash-Harsanyi es por completo independiente del proceso seriado que ese enfoque pretende establecer. En nuestra perspectiva, aunque sólo haya una pretensión y una concesión de parte de cada negociador, la pretensión importa y por lo tanto la concesión se basa en algo que importa. El proceso no es algo fingido. Y esto, insistimos, capta más adecuadamente la naturaleza de la negociación. Suponemos con seguridad que la pretensión inicial de cada negociador debe ser significativa y que toda la concesión que haga de esa pretensión debe ser de interés central. Nuestro enfoque de la negociación satisface esas suposiciones. Y representa a personas racionales que realmente están negociando.

4.1 Durante todo este largo análisis de la racionalidad interna de la cooperación la justicia ha estado ausente. Pero, como veremos inmediatamente, la justicia y la razón coinciden en un ideal único de interacción cooperativa. El principio de concesión relativa minimáxima no sólo sirve como base de un acuerdo racional sino también como fundamento de una restricción imparcial a la conducta de cada individuo. Y la justicia es la disposición sujeta a esa restricción.

Al considerar que la interacción cooperativa forma parte del dominio de la justicia hacemos una afirmación doble. La primera es que, como el mercado, la cooperación racional excluye toda parcialidad. La segunda es que, a diferencia del mercado, esa exclusión exige que cada cooperador restrinja su actividad maximizadora. La interacción que logra la imparcialidad sin restricción constituye una zona moralmente libre, de la cual están ausentes las externalidades que conducen a los maximizadores de la utilidad a convertirse en jugadores líberos y en parásitos. Pero la interacción cooperativa tiene que vérselas con esas externalidades; la cooperación es la mano visible que restringe a las personas y les impide tomar ventaja de sus pares, pero que las restringe imparcialmente y de un modo beneficioso para todos. Tal restricción atrae la aceptación racional; y ésa es la idea que subyace en una moral basada en el acuerdo.

La argumentación que desarrollaremos en este capítulo quedará limitada de dos maneras por completo diferentes. Primero, puesto que sólo examinamos el proceso de negociación y dejamos de lado la posición negociadora inicial, consideraremos sólo la imparcialidad o la justicia de ese proceso. Dada una posición inicial, la cooperación es justa si la estrategia conjunta en la que se basa es el resultado de una negociación justa entre los cooperadores. Pero la justicia del proceso negociador no corrige cualquier parcialidad que pueda estar presente en la posición inicial; en realidad, sólo podría transmitir la parcialidad de la posición inicial a la estrategia conjunta seleccionada. En el capítulo 7 examinaremos las condiciones necesarias para que la posición negociadora inicial sea imparcial.

En segundo lugar, abordaremos la imparcialidad como abordamos la racionalidad, desde el punto de vista del individuo realmente implicado en la negociación. Pero, en el capítulo 1 ya mencionamos el hecho de que otros teóricos intentaron vincular la razón con la moral abordando la imparcialidad en una perspectiva aparentemente diferente, es decir, considerando qué principios de la interacción debería elegir cualquier individuo que ocupe un punto de vista imparcial especialmente concebido. Aunque en el capítulo 8 mostraremos que esa perspectiva armoniza con la nuestra, esto no es lo que nos ocupa en

este momento. Ahora nos concentraremos en la elección cooperativa de una estrategia conjunta, elección que es imparcial porque toda persona implicada, situada en cualquier perspectiva, la considera aceptable.

Obsérvese que en esta elección no está incluida ninguna restricción a la conducta maximizadora. La negociación es una actividad directamente maximizadora que conduce a un acuerdo sobre una estrategia conjunta. La restricción forma parte de la interacción cooperativa, que exige la adhesión a esa estrategia aun cuando en general el resultado no esté en equilibrio. En esta sección nos interesa mostrar que si la cooperación se da partiendo de un acuerdo racional, la restricción que imponga esa cooperación ha de ser justa. Pero alguien aun puede preguntarse, con el Necio de Hobbes, si es racional ser justo, adherirse a la restricción que uno ha acordado racionalmente.[18] Nos referiremos a esta pregunta sobre el acatamiento o la adhesión en el capítulo 6.

4.2 Recordemos que la cooperación aparece como un intento de alcanzar un resultado óptimo en situaciones en las cuales la presencia de externalidades hace que el resultado de la interacción natural o de la interacción del mercado sea subóptimo. La significación moral de las externalidades se basa en la posibilidad de que una persona pueda sacar ventaja de otra, ya sea como jugador líbero, obteniendo algún beneficio exento de costo, derivado de las actividades de los demás, ya sea como parásito, transfiriéndoles a los demás el costo de algún beneficio que recibe. A fin de evitar esa posibilidad, la cooperación debe asegurar que la razón o proporción entre el beneficio que recibe el cooperador y la contribución que haga sea, en la medida de lo posible, constante y la misma para todos. Pero, ¿cómo se logra esto?

Comencemos con un ejemplo sencillo en el cual supondremos que los frutos de la cooperación —el superávit cooperativo, cuya distribución está en juego en el proceso negociador— pueda considerarse como un bien único transferible.[19] La cantidad de ese bien producido mediante la interacción cooperativa es fija; puede distribuirse de cualquier modo entre los cooperadores sin que por ello se afecte su cantidad. Las utilidades que obtiene cada persona deben ser lineales respecto de la parte de ese bien

que le corresponde, pero no suponemos que puedan hacerse comparaciones interpersonales de las utilidades. En nuestro ejemplo supondremos que ese bien es dinero.

Nuestro ejemplo es la sociedad constituida por Abel y Mabel que mencionamos al final de 3.2. Abel y Mabel esperan obtener una retribución k de la inversión conjunta. Pero k no es el superávit cooperativo. Ese superávit es k menos la suma de la retribución que cada uno podría esperar de la inversión independiente, que es en el caso de Abel cr y en el caso de Mabel $(k - 2kr + cr)$. De modo que el superávit cooperativo es $(2kr - 2cr)$.

¿Cómo se divide ese superávit entre Abel y Mabel? Abel, como mostramos en 3.2, recibe kr de la cooperación; si sustraemos su rendimiento inicial cr, comprobamos que la parte que le corresponde del superávit cooperativo es $(kr - cr)$ o, exactamente 1/2. Mabel recibe $k(1 - r)$ de la cooperación; si sustraemos su rendimiento inicial $(k - 2kr + cr)$ comprobamos que la parte que le corresponde del superávit cooperativo es, por supuesto, también $(kr - cr)$. Si cada uno recibe un rendimiento total proporcional a su contribución, cada uno recibe pues la mitad del superávit cooperativo. Puesto que suponemos que las retribuciones se dividen entre Abel y Mabel sobre la base de sus contribuciones, esto muestra que ambos contribuyen igualmente a producir el superávit cooperativo.

Nuestro análisis puso de manifiesto un resultado que seguramente puede parecer sorprendente. Si Abel y Mabel han de comprometerse en una sociedad en la cual sus resultados han de ser proporcionales a sus contribuciones, deben pues dividir el superávit cooperativo logrado por su asociación en partes iguales. Una división equitativa del superávit puede no parecer sorprendente en sí misma, porque puesto que ninguno de los dos puede ganar una parte de ese superávit sin el otro, ambos son igualmente responsables de que tal superávit esté disponible, y por consiguiente, cada uno tiene derecho a una parte igual. Lo que resulta sorprendente es que esa visión igualitaria de la cooperación se manifieste plenamente compatible con la exigencia razonable de que las retribuciones generales sean proporcionales a las (diferentes) contribuciones generales.

En realidad, esta visión igualitaria de la cooperación está implícita en nuestro análisis. El superávit cooperativo es en el

sentido más acabado el producto conjunto de los cooperadores. Nadie puede razonable o justamente pretender más, ni nadie debería razonable o justamente aceptar menos que una distribución igualitaria del superávit cooperativo, en el caso de que pudieran determinarse esas partes iguales. Si hay un único bien transferible, producido en una cantidad fija y divisible de cualquier forma entre los cooperadores, la racionalidad y la imparcialidad exigen pues esa división equitativa. Es evidente que ese es el resultado que exige nuestra teoría de la negociación porque, dado un bien fijo, completamente divisible, cada individuo pretende obtenerlo todo y la concesión máxima queda pues minimizada si y sólo si cada individuo recibe una parte igual del bien.

Pero podría objetarse que semejante resultado es absurdo. La señora Macquarrie, química farmacéutica, necesita un ayudante de laboratorio que la ayude a realizar sus experimentos y contrata al señor O'Rourke. Cuando de resultas de sus experimentos la señora Macquarrie descubre una droga maravillosa que puede hacerla millonaria, ¿debe dividir sus derechos de invención con el señor O'Rourke? Evidentemente no. Ella no realizó sus experimentos como una empresa conjunta con su asistente. Aunque necesitaba un asistente, no necesitaba precisamente al señor O'Rourke.Su relación con el señor O'Rourke fue estrictamente una transacción de mercado; la señora Macquarrie lo contrató por (presumiblemente) el porcentaje de ganancias que le corresponde a los asistentes de laboratorio. Quizá podríamos decir que sería una persona poco generosa si no lo gratificara con una buena recompensa, pero evidentemente no le debe nada.

Consideremos, por otra parte, el aprieto en que se encuentra Sam McGee, un prospector de minas que descubre la vena de oro más rica del Yukon, pero que carece del dinero necesario (digamos $100) para registrar su derecho a ella.[20] Si Grasp, el banquero, es el único hombre de Dawson City que tiene el dinero disponible para proporcionárselo a McGee, el pobre Sam (racionalmente) tiene que ofrecerle a Grasp la mitad de los derechos. Porque si bien los $100 de Grasp sólo valen $100 en ausencia del descubrimiento de McGee, el descubrimiento carece de valor sin el dinero de Grasp. Por supuesto que si hubiera otras fuentes de

dinero, Sam se encontraría en una posición comparable a la de la señora Macquarrie; necesitaría el dinero, pero no el dinero de Grasp y en ese caso podría obtenerlo al interés habitual.

4.3 ¿Qué ocurre cuando lo que hay no es un bien único transferible, producido en una cantidad fija y divisible a voluntad entre los cooperadores? ¿Cómo podemos determinar la división justa de las partes de un superávit cooperativo? Recordemos la significación de la pretensión que presenta cada persona. Cada individuo pretende lo máximo posible que puede recibir del superávit, pero sólo de esas interacciones cooperativas de las cuales participa. Por lo tanto, si no se hace ninguna contribución no se pretende nada; pero si se hace alguna contribución se pretende lo máximo. Una pretensión máxima es pretender todo el superávit, en la medida en que el que lo pretenda pueda recibirlo. Igualemos entonces las pretensiones máximas de los negociadores: cada uno presenta una pretensión equivalente a la de los demás. Supongamos que todos reciben la misma proporción de sus pretensiones. Entonces, puesto que sus pretensiones son equivalentes, todos reciben partes iguales del superávit cooperativo.

En la subsección anterior dijimos que cada persona que contribuye a la interacción cooperativa es igualmente responsable del superávit resultante. Podemos modificar esa afirmación si la cooperación no da por resultado una cantidad fija de un bien completamente transferible y decir que cada persona tiene una responsabilidad equivalente a aquello que puede recibir del superávit cooperativo. De modo que cada uno tiene un derecho igual al superávit en la medida en que los productos de la cooperación estén disponibles para él. Aunque no tengamos una medida que nos permita dividir el superávit en partes iguales, (en general) podemos dividirlo en partes que son equivalentes desde el punto de vista de las pretensiones máximas y que por lo tanto son derechos adquiridos por los negociadores. Por consiguiente, si una distribución justa o imparcial del superávit cooperativo relaciona el beneficio que recibe cada persona con la contribución que ésta hace, la parte que le corresponde justamente a cada persona se determina estableciendo partes proporcionales a las pretensiones.

Ahora estamos en posición de establecer un principio de

negociación justa. Supongamos, como en 3.2 que la posición negociadora inicial le proporciona a una persona una utilidad u^* y que esa persona pretende un resultado que le proporcione una utilidad u^*. De modo que su pretensión respecto del superávit cooperativo es $(u^* - u^*)$. Supongamos que la estrategia conjunta elegida para la interacción cooperativa le proporcione una utilidad esperada u; por lo tanto la parte esperada que le corresponde del superávit cooperativo es $(u - u^*)$. La proporción de su pretensión referente al superávit cooperativo que recibe es pues $[(u - u^*)/(u^* - u^*)]$; llamaremos a esto su *beneficio relativo.* Obsérvese que el beneficio relativo debe caer entre 0 —que es el beneficio relativo de la posición negociadora inicial $[(u^* - u^*)/(u^* - u^*)]$— y 1 —que es el beneficio relativo de la pretensión $[(u^* - u^*)/(u^* - u^*)]$—. Si tenemos que dividir las partes de manera proporcional a las pretensiones, debemos pues dividir el superávit cooperativo de modo tal que cada persona reciba el beneficio relativo máximo igual.

Esto puede no ser siempre posible. O, antes bien, puede ocurrir que si cada persona recibe el beneficio relativo máximo igual, no todo el superávit cooperativo se distribuya entre los cooperadores. En ese caso, parecería razonable maximizar el beneficio relativo mínimo. Puesto que esto equivale a maximizar el beneficio relativo igual cuando el último consume el superávit cooperativo, podemos decir que una distribución justa o imparcial del superávit cooperativo —y por lo tanto una elección justa o imparcial de la estrategia conjunta para la interacción cooperativa— debe estar de acuerdo con el *Principio de maximínimo* (máximo-mínimo) *beneficio relativo.*

Pero ésta es nuestra vieja amiga, la concesión relativa minimáxima, con un nuevo ropaje. Porque el beneficio relativo y la concesión relativa sumados son la unidad. $[(u - u^*)/(u^* - u^*)] + [(u^* - u)/(u^* - u^*)]=1$. Un individuo maximiza el beneficio relativo mínimo minimizando la concesión relativa máxima. En la negociación la imparcialidad y la racionalidad coinciden.

Este resultado puede no parecernos sorprendente. Porque aquellas consideraciones que justificadamente pudieron llevarnos a cuestionar la justicia de muchas instancias de la negociación cotidiana no tienen lugar en el contexto de la racionalidad plena.En las negociaciones corrientes, las personas pueden

esconder rasgos significativos de sus circunstancias. , o todas las opciones de que disponen pueden representar mal sus preferencias o la intensidad de sus preferencias. Pero suponemos que cada persona tiene una información completa, que conoce las acciones que pueden realizar todas las demás, los resultados que pueden surgir de tales acciones y los rendimientos de utilidad que cada uno de esos resultados ha de proporcionarle a todos los negociadores. En las negociaciones corrientes las personas pueden simular, especialmente si pueden ocultar o representar erróneamente algunos factores, de modo tal que los demás se hacen expectativas inciertas o erradas sobre lo que quienes simulan están dispuestos a hacer. Pero aquí no hay lugar para la simulación; no sólo porque cada individuo está plenamente informado sino porque además es un maximizador de la utilidad racional que sabe que sus pares también lo son. En la negociación corriente las personas pueden amenazar, pero entre las personas plenamente racionales las amenazas son inútiles; nadie le creería a otro que declara que ha de obrar de un modo no maximizador de su utilidad si los demás se amedrentan ante su amenaza, y decir que uno ha de obrar de un modo maximizador de su utilidad no constituye una amenaza.[21] Nuestros negociadores no tienen fuerzas psicológicas que les permitan explotar a los demás, ni debilidades psicológicas que dejen que los demás los exploten. Y partimos de la base de que la negociación está libre de costos, tanto desde el punto de vista de la utilidad como del tiempo, de modo tal que nadie tiene necesidad de llegar a una conclusión sin considerarlo antes plenamente, la negociación no está bajo presión. Por lo tanto, cada negociador puede utilizar su propia racionalidad para apelar a la misma racionalidad de sus semejantes. Además de la racionalidad, lo único que hay que considerar son las preferencias y las posibles acciones de cada persona, y sobre esa base negocian todos los individuos.

Podríamos suponer que aun dando por sentada la naturaleza ideal de nuestros negociadores, las circunstancias podrían sin embargo dar lugar a una parcialidad significativa. Pero los problemas referentes a las circunstancias corresponden a la posición negociadora inicial y no al proceso de negociación. En el capítulo 7 nos referiremos a las circunstancias. Ciertamente,

si el ideal de la interacción cooperativa es eliminar el juego líbero y el parasitismo característicos de la interacción natural, haría pues falta aplicar cierta restricción al identificar la posición inicial con el resultado no cooperativo, puesto que ese resultado puede incorporar los efectos de las actividades reales que la cooperación eliminaría. Pero aquí lo que nos ocupa es la elección como base para la interacción cooperativa (dando por sentada una posición inicial) y lo que sostenemos es que al proporcionar partes iguales del superávit cooperativo a todas las personas, el principio de concesión relativa minimáxima garantiza que la imparcialidad de la negociación vincule la contribución que hace cada persona con la cooperación, con el beneficio que recibe de ella.

Concluiremos esta discusión haciendo notar que muchos de nuestros principios y prácticas morales verdaderos son en realidad aplicaciones de las exigencias de la concesión relativa minimáxima a contextos particulares. Podemos suponer que principios tales como cumplir las promesas, decir la verdad o negociar honestamente han de defenderse mostrando que la adhesión a ellos les permite a las personas cooperar de un modo que, es lógico esperar, equipare, por lo menos de una manera general, los beneficios relativos que proporciona la interacción. Estas están entre las prácticas fundamentales de la moral que podemos recomendar a cada individuo, mostrándole que la moral rige su acuerdo racional.

Notas

1 J. Rawls, *A Theory of Justice,* Cambridge, Mass. 1971, pág. 126.

2 Hume, *Treatise,* iii,ii,ii, págs. 486-488; véase también *Enquiry,* iii,i, págs. 183-186.

3 Véase Hobbes, *Leviathan,* cap. 13, págs. 60-62.

4 *Ibid,* cap. 13, pág.63.

5 Véase Ruth Benedict, *Patterns of Culture,* Boston y Nueva York, 1934, págs. 139-140, 146-148.

6 Véase K. J. Arrow, "Values and Collective Decision-Making", en P. Laslett y W. G. Runciman (comps.), *Philosophy, Politics and Society,* 3 series, Oxford, 1967, págs. 225-226.

7 Véase R. D. Luce y H. Raiffa, *Games and Decisions,* Nueva York, 1957, págs. 344-345.

8 Véase la discusión de A. K. Sen, *Collective Choice and Social Welfare,* San Francisco, 1970, págs. 47-50. Nuestra afirmación es más fuerte que cualquiera de las defendidas por Sen y se basa en un trabajo en preparación.

9 La prueba es sencilla y se demuestra mediante ejemplos. Consideremos una situación con tres resultados óptimos posibles representados por los puntos (0,1, 1), (0,7, 0,7) y (1, 0,1) y un cuarto resultado posible representado por (0, 0). Una lotería que le asigne las mismas probabilidades a cada uno de los resultados óptimos determina el resultado esperado representado por el punto (0,6, 0,6) que, evidentemente, es subóptimo.

10 Aunque la terminología no sea la suya el punto de vista sí es el de Harsanyi. Harsanyi sostiene que "la versión actualizada del utilitarismo clásico es la única teoría ética que cumple consistentemente con el principio que establece que las cuestiones morales deben decidirse mediante tests racionales y que la conducta moral misma es una forma especial de la conducta racional", "Morality and the theory of rational behaviour", pág. 40.

11 Puede hallarse la demostración en J. C. Harsanyi, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", en *Essays on Ethics,* págs. 10-15; también en *Rational Behavior and Bargaining Equilibrium in Games and Social Situations,* Cambridge, 1977, págs. 64-69.

12 Véase A. E. Roth, M. W. K. Malouf y J. K. Murnighan, "Sociological versus Strategic Factors in Bargaining", *Journal of Economic Behavior and Organization* 2, 1981, págs. 174-177.

13 Véase F. Zeuthen, *Problems of Monopoly and Economic Welfare,* Londres, 1930, cap.4; J. F. Nash, "The Bargaining Problem", *Econometrica* 18, 1951, págs. 155-162 y "Two-person Cooperative Games", *Econometrica* 21, 1953, págs. 128-140; Harsanyi, *Rational Behavior.*

14 La idea en la que se basa nuestra teoría aparece en "Rational Co-operation", *Noûs* 8, 1974, págs. 53-65. Aparece también un breve esbozo en "Economic Rationality and Moral Constraints", *Midwest Studies in Philosophy* 3, 1978, págs. 92-93. y un análisis más completo en "The Social Contract: Individual Decision or Collective Bargain?", en C.A. Hooker, J. J. Leach y E. F. McClennen (comps.), *Foundations and Applications of Decision Theory,* 2 vols., Dordrecht, 1978, vol. 2, págs. 47-67. También puede encontrarse una discusión informal en "Bargaining Our Way Into Morality: A Do-It-Yourself Primer", *Philosophic Exchange* 2, n°5, 1979, págs. 14-27. Véase también el más reciente "Justified Inequality?", en *Dialogue* 21, 1982, págs. 431-443 y "Justice as Social Choice", en D. Copp y D.

Zimmerman (comps.), *Morality, Reason and Truth: New Essays on the Foundations of Ethics*, Totowa, NJ, 1985, págs. 251-269. En las situaciones que de las que participan sólo dos personas, nuestro enfoque brinda una solución *G*, analizada en A. E. Roth, *Axiomatic Models of Bargaining*, Berlín, 1979, págs. 98-108, y axiomatizada por E. Kalai y M. Smorodinsky, "Other Solutions to Nash' Bargaining Problem", *Econometrica* 43, 1975, págs. 513-518. Pero en las situaciones en que participan más de dos personas, nuestra explicación se aparta de la solución *G* en aquellos casos en que *G* no puede ser Pareto-óptima. Véase mi artículo "Bargaining and Justice", aparecido en *Social Philosophy and Policy* 2, n°2.

15 Roth, *op. cit.*, págs. 105-107.

16 Véase además de las referencias a Nash y Zeuthen mencionadas en la nota 13 *supra*, Luce y Raiffa, págs. 124-137. Harsanyi hace un breve análisis del trabajo anterior en *Rational Behavior*, págs. 143-153.

17 Véase Harsanyi, *Rational Behavior*, págs. 118, 153-162.

18 Véase Hobbes, *Leviathan*, cap. 15, págs. 72-73.

19 Sobre una explicación de lo que es un bien transferible, véase Luce y Raiffa, págs. 168-169.

20 Sobre S. McGee, véase "The Cremation of Sam McGee", en R. W. Service, *Songs of a Sourdough*, Toronto, 1907.

21 En el capítulo 6 sección 3.2 sostendremos que las amenazas son compatibles con la racionalidad, pero no en el contexto de la cooperación. En efecto, la conducta amenazadora debería proscribirse mediante una restricción sobre la posición negociadora inicial, que desarrollaremos en el capítulo 7.

VI

El acatamiento: un modo de restringir la maximización

1.1 La persona justa está dispuesta a acatar las exigencias del principio de concesión relativa minimáxima, al interactuar con aquellos semejantes que, según cree, están igualmente dispuestos. La persona justa se adapta a la sociedad pues ha internalizado la idea del beneficio mutuo, de modo tal que al elegir su curso de acción dedica una consideración primaria a la perspectiva de alcanzar un resultado cooperativo. Si esa persona tiene la posibilidad de lograr, o puede razonablemente esperar alcanzar, un resultado que sea tanto justo (aproximadamente) como óptimo (aproximadamente), elige obrar del modo que le permita alcanzarlo; sólo si no puede razonablemente esperar que aquello que elija haya de maximizar su propia utilidad.

A fin de vincular nuestra caracterización de la persona cooperativa con las condiciones que debe cumplir la interacción racional establecidas en el capítulo 3, definiremos una estrategia optimizadora justa (o una elección, o una respuesta) como aquella que, teniendo en cuenta las estrategias esperadas de los demás, le permita esperar al individuo un resultado (aproximadamente) justo y óptimo, un resultado cuyos rendimientos de utilidad sean similares a los del resultado cooperativo, como lo determina la concesión relativa minimáxima. Decimos que la respuesta es aproximadamente justa y óptima porque en muchas situaciones una persona no ha de esperar que los demás hagan precisamente lo que les exigiría la concesión relativa minimáxima, de modo tal que esa persona no puede elegir una estrategia cuyo resultado sea completamente justo o plenamente óptimo. Pero suponemos que aun así esa persona sigue estando dispuesta a cooperar antes que a elegir la interacción no cooperativa.

De modo que una persona justa acepta esta formulación de la condición *A*: *A'*: La elección de cada persona debe constituir una respuesta optimizadora justa a la elección que espera

hagan los demás, siempre que pueda optar por esa respuesta; de lo contrario, su elección debe constituir una respuesta maximizadora de la utilidad. Una persona justa está dispuesta a interactuar con los demás sobre la base de la condición *A'*.

Sin embargo, una persona justa debe cobrar conciencia de que nos todas las personas racionales aceptan esa formulación de la condición original *A*. Al formarse expectativas referentes a las elecciones de los demás, el individuo no tiene por qué suponer que las elecciones de esos otros hayan de satisfacer la condición *A'*. Por consiguiente, no podemos renunciar a las condiciones originales *A*, *B* y *C*, como condiciones necesarias para la interacción estratégica; la "respuesta racional" permanece abierta a varias interpretaciones (por lo menos hasta que nuestra teoría alcance la aceptación universal).

En este capítulo la tarea que nos hemos propuesto cumplir es darle a la condición *A'* un carácter racional maximizador de la utilidad. Y lo haremos demostrando que, dadas ciertas condiciones plausibles y razonables, un maximizador racional de la utilidad, que debe elegir entre no aceptar ninguna restricción a las elecciones que haga en la interacción y aceptar las restricciones a sus elecciones que le exige la concesión relativa minimáxima, elige esta última opción. La persona elige cómo hacer otras elecciones adicionales; elige en una perspectiva maximizadora de la utilidad no continuar haciendo elecciones en esa perspectiva.

Al defender la condición *A'*, estamos defendiendo el consentimiento o la aceptación de acuerdos basados en el principio de la concesión relativa minimáxima. En realidad, estamos defendiendo el consentimiento no sólo referido a los acuerdos sino a las prácticas que concordarían con ese principio o estarían apoyadas en él. Si nuestra defensa fracasa, debemos pues llegar a la conclusión de que la negociación racional es vana y de que la cooperación, aunque se base en acuerdos racionales, no constituye en sí misma una exigencia racional, de modo tal que no nos permite superar las deficiencias de la interacción natural y de la interacción del mercado. En realidad, si nuestra defensa fracasa, debemos pues llegar a la conclusión de que la moral racional es una quimera y que por lo tanto no existen restricciones imparciales y racionales a la búsqueda de la utilidad individual.

Al defender la condición *A'*, estamos defendiendo la racionalidad externa de la cooperación contra las objeciones del egoísta. Sea lo que fuere que tenga posibilidad de hacer, el egoísta siempre busca maximizar su utilidad esperada. Aun reconociendo que la cooperación le ofrece la perspectiva del beneficio mutuo, el egoísta niega que comportarse de manera cooperativa sea racional, si esa conducta restringe la maximización. Ese egoísta hace su aparición filosófica en el Necio de Thomas Hobbes, en el *Leviathan*, como lo mostraremos seguidamente.

1.2 Hobbes comienza a desarrollar su teoría moral con una concepción puramente permisiva del derecho de la naturaleza, estableciendo lo que un individuo puede hacer y no lo que debería permitírsele hacer o lo que debería hacerse por él. El permiso es racional porque, como dice Hobbes, "la palabra *derecho* no significa ninguna otra cosa más que la libertad que tiene todo hombre para utilizar sus facultades naturales de acuerdo con la recta razón".[1] Y Hobbes afirma que en la condición natural de la humanidad esa libertad es ilimitada, de modo tal que "todo hombre tiene un Derecho a todas las cosas, incluso al cuerpo de los demás". Al concebir de ese modo el derecho natural, Hobbes expresa una visión directamente maximizadora de la acción racional, sujeta a la condición natural, central para su psicología, que sostiene que cada individuo busca ante todo su propia preservación. Para Hobbes, toda persona tiene el derecho inicial de hacer lo que esté a su alcance para preservarse, pero los demás no tienen ninguna obligación ni de dejarlo hacer ni de hacer por él lo que sea necesario para preservarlo.

Hobbes sostiene que en la condición en la que todas las personas ejercen ese derecho ilimitado, "ningún hombre (por fuerte e inteligente que sea) puede tener la seguridad de vivir más tiempo del que la Naturaleza normalmente les permite vivir a los hombres".[3] Las personas que buscan su propia preservación se encuentran trabadas en un combate mortal. Pero si bien la razón pone a los hombres en esa condición de guerra, también puede sacarlos de ella. Hobbes dice "la razón indica los convenientes Artículos de la Paz, basándose en los cuales los hombres pueden llegar a un acuerdo. Esos Artículos

son lo que llamamos las Leyes de la Naturaleza".[4] Las leyes de la naturaleza son preceptos, "descubiertos por la Razón, según los cuales a un hombre le está prohibido hacer aquello que es destructivo para su vida o que le impide tener acceso a los medios de preservarla y le está prohibido omitir aquello que considera que es lo que mejor lo protege".[5]

Puesto que la guerra es enemiga de la preservación, la primera ley o la ley fundamental de la naturaleza es "que cada hombre se esfuerce por lograr la Paz, en la medida en que tenga la esperanza de alcanzarla", a lo que Hobbes agrega, "y cuando no pueda lograrla lo que debe buscar y utilizar son todas las ayudas y ventajas que pueda proporcionarle la Guerra".[6] De esto Hobbes extrae inmediatamente otra ley, que establece como medio fundamental para lograr la paz, "que cada hombre esté dispuesto, cuando los demás también lo estén y en la medida que lo considere necesario para alcanzar la paz y para defenderse a sí mismo, a renunciar a su derecho a todas las cosas; y que se contente con tener tanta libertad ante los demás como la que está dispuesto a permitirle a los demás tener ante él".[7] Puesto que el derecho ilimitado de la naturaleza da lugar a la guerra, renunciar a cierta parte de ese derecho es algo necesario para alcanzar la paz. Por supuesto, la renuncia debe ser mutua; cada persona espera beneficiarse no gracias a su propia renuncia, sino gracias a la renuncia de sus semejantes, de modo tal que ninguno de ellos tiene razones para renunciar a sus derechos unilateralmente. Lo que imagina Hobbes es una negociación racional en la cual cada individuo acepta ciertas restricciones a su libertad de acción, de modo tal que todos puedan evitar los costos de la condición natural de la guerra.

La defensa de esta segunda ley es directa. Hobbes sólo necesita decir que "mientras todos los hombres conserven ese derecho a hacer lo que mejor les plazca, todos los hombres estarán en la condición de la Guerra".[8] Del mismo modo sencillo defiende la reciprocidad que exige la ley: "si los demás hombres no renuncian a sus Derechos, como lo hace él; ninguno tiene pues razones para enajenar su Derecho. Porque eso sería exponerse al Robo (a lo que ningún hombre está obligado) en lugar de disponerse a la Paz".[9] Para cada individuo es directamente ventajoso acordar con sus semejantes la mutua renuncia al

derecho y por lo tanto la mutua aceptación de la restricción. Hobbes concibe esa restricción como una obligación, a la que sólo se puede llegar mediante el acuerdo, puesto que "sobre ningún hombre puede pesar una obligación que no nazca de alguna decisión suya, porque todos los hombres son igualmente libres por Naturaleza".[10] La teoría de Hobbes, como la nuestra defiende una moral basada en el acuerdo.

Hobbes reconoce que una cosa es hacer un acuerdo o un pacto y otra muy diferente cumplirlo. Hobbes no supone que la segunda ley de la naturaleza, al ordenarnos llegar a un acuerdo, también nos ordena acatarlo. Por lo tanto presenta una tercera ley de la naturaleza: "que los hombres realicen lo que establecen sus Pactos" que constituye para Hobbes "la Raíz de la JUSTICIA".[11] Una persona justa es aquella que cumple los acuerdos que estableció racionalmente.

La defensa que hace Hobbes de esta tercera ley ya no es tan directa como la de la segunda. Como él mismo reconoce, sin ella "los Pactos son vanos y no son sino palabras Vacías; y el Derecho que tienen todos los hombres al resto de las cosas continúa poniéndonos en la condición de la Guerra."[12] Pero esto no muestra que aceptarla proporcione ningún beneficio directo. Al hacer un pacto, cada persona maximiza su utilidad esperada, puesto que cada uno obtiene una ganancia de la renuncia recíproca que ese pacto implica. Pero cada individuo no maximiza su utilidad esperada al establecer un pacto, en la medida en que ese pacto le exige limitar el ejercicio de alguna parte de su libertad previa. Y esto abre la puerta a la objeción del Necio.

Dejemos que Hobbes hable por sí mismo.

> El Necio se dijo íntimamente que no existe tal cosa llamada Justicia; y a veces hasta lo expresó con palabras, afirmando seriamente que estando la conservación y la satisfacción de cada hombre a su propio cuidado, no habría razón para que cada uno no hiciera aquello que piensa que ha de conducirlo a lograrlas; y por consiguiente, no sería contrario a la Razón establecer o no establecer, acatar o no acatar, Pactos si éstos han de aumentar el propio beneficio. De ahí que él no niegue la conveniencia de que se establezcan pactos, que a veces se violan y que a veces se cumplen y que considere que tales violaciones puedan llamarse Injusticia y que la observancia de esos pactos se llame Justicia: pero el Necio se pregunta si la Injusticia... a veces no puede estar

en armonía con esa Razón que le dicta a cada hombre que busque su propio bien...[13]

El Necio no se opone seriamente a la segunda ley de la naturaleza porque Hobbes supone que cada individuo sólo ha de hacer aquellos pactos que considere ventajosos y ésa no es una conducta que objete el Necio. Lo que sí objeta es la tercera ley, la ley que exige el consentimiento o la adhesión a los pactos establecidos por cada individuo porque, por ventajoso que pueda resultar hacer un pacto, ¿no puede ser aun más ventajoso violarlo? Y si es ventajoso, ¿no es entonces racional? El Necio desconfía del núcleo mismo de la conexión entre la razón y la moral que tanto Hobbes como nosotros pretendemos establecer: la racionalidad de aceptar una restricción moral a la búsqueda directa de la mayor utilidad individual.

1.3 Al replicarle al Necio, Hoobes afirma que, si existe la suficiente seguridad de que una de las partes ha de acatar el pacto, la pregunta es "si es contrario a la razón, es decir, si es contrario al beneficio del otro, acatar o no el pacto".[14] La interpretación más natural indica que Hobbes pregunta si es racional acatar el propio pacto, lo cual constituye una respuesta maximizadora de la utilidad al acatamiento del pacto hecho por los pares. Si éste fuera en realidad el punto de vista de Hobbes, podríamos decir que Hobbes se esfuerza por refutar al Necio apelando, en efecto, a la condición *A* de la elección estratégicamente racional, es decir, considerando que una respuesta racional es simplemente una respuesta maximizadora de la utilidad. Con lo cual no podríamos abrigar muchas esperanzas de que Hobbes tuviera éxito en esa discusión.

El primer argumento esgrimido por Hobbes le recuerda al Necio que la racionalidad de la elección depende de las expectativas y no de los resultados reales. Esto no debe necesariamente detenernos. Su segundo argumento aborda la cuestión en un nivel más profundo.

> El que...viole su Pacto y consecuentemente declare que puede hacerlo con razón, no puede ser recibido en el seno de ninguna Sociedad que una a las personas para la Paz y la Defensa, pero si por error esa Sociedad lo recibe y lo retiene en su seno, no puede dejar de advertir el peligro que entraña tal error, error

> que un hombre no puede razonablemente considerar como medio de su seguridad.[15]

Una persona dispuesta a violar sus pactos no puede ser aceptada como parte de los acuerdos cooperativos por aquellos individuos que no sólo son racionales sino que además tienen conciencia de la mala disposición del primero, por consiguiente, una persona no puede razonablemente esperar sacar provecho de los beneficios que están disponibles para los cooperadores. Aun cuando sus violaciones particulares del pacto pudieran beneficiarlo, la disposición que lo lleva a realizar tales violaciones no.

En efecto, Hobbes traslada la cuestión desde preguntarse si acatar los propios acuerdos es contrario a la razón entendida ésta como maximización de la utilidad (suponiendo que haya la suficiente seguridad de que los demás acatan sus propios acuerdos) a preguntarse si es contrario a la razón estar dispuesto a acatar los propios acuerdos. La disposición a decidir si uno ha de adherirse o no a los propios acuerdos o pactos apelando a consideraciones directamente maximizadoras de la utilidad es en sí misma desventajosa si se conoce o si se sospecha lo suficiente, porque lo excluye a uno de la participación con aquellos que sospechan tal disposición en aquellos arreglos cooperativos en los que lograr los beneficios exige que cada participante renuncie a maximizar su utilidad o para decirlo con la terminología de Hobbes, exige que cada individuo renuncie a cierta porción de su derecho natural original e ilimitado. La disposición a acatar los propios acuerdos, suponiendo que exista la seguridad suficiente, sin apelar a consideraciones directamente maximizadoras de la utilidad, convierte al individuo en un socio deseable para la cooperación beneficiosa y por lo tanto es en sí misma beneficiosa. Probaremos más adelante que ésta es la clave para demostrar que un maximizador de la utilidad plenamente racional se dispone a acatar sus pactos o acuerdos racionalmente convenidos.

Pero para que Hobbes obtenga una ventaja plena de su respuesta al Necio, es necesario que revise su concepción de la racionalidad y quiebre la conexión directa entre la razón y el beneficio con la que comienza su réplica. Hobbes tiene que decir

que es racional acatar el propio pacto aun cuando el acatamiento no le proporcione un beneficio directo, siempre que sea beneficioso para uno estar dispuesto a acatar el acuerdo. Pero Hobbes nunca dice eso. Y en la medida en que permite que el Necio vincule directamente la razón con el beneficio que implica el acatamiento, antes que vincularlo con el beneficio que implica la disposición al acatamiento, también le permite rehuir la refutación.

En su discusión con el obispo Bramhall, Hobbes sugiere una revisión de su concepción de la racionalidad. Al estar de acuerdo con Bramhall en cuanto a que "la bondad moral es la concordancia de una acción con la correcta razón", no declara que aquello que es moralmente bueno conduce al propio beneficio, sino que sostiene en cambio que:

> Todo lo realmente bueno...es aquello que no repugna a la ley...pues la ley representa toda la razón recta que tenemos, y... es la norma infalible de la bondad moral. La razón, sea de quien fuere, es esto, porque ni la razón mía ni la del obispo llegan a ser una norma de nuestras acciones morales, por lo tanto hemos puesto sobre nosotros un gobernador soberano y hemos acordado que sus leyes ...han de dictarnos lo que es realmente bueno.[16]

Al argumento presentado por el Necio sosteniendo que la injusticia puede "a veces puede estar en armonía con esa Razón que le dicta a cada hombre que busque su propio bien",[17] Hobbes puede responder que la injusticia no puede apoyarse en esa razón que está constituida por la ley del soberano. Así como no le conviene a cada hombre conservar su derecho natural completo, tampoco le conviene conservar su razón natural como guía para sus acciones. Pero Hobbes no supone que cada hombre internaliza la correcta razón del soberano. Su psicología egoísta permite únicamente la internalización de una norma que se interese directamente por la preservación y la satisfacción individual. Y esto sólo es así en la medida en que el soberano pueda poner en vigor la ley a cuyo acatamiento esté racionalmente obligado el individuo. Pero esto es proponerle una solución política y no moral al problema planteado por el Necio.

Si el mercado obra como una mano invisible y dirige los esfuerzos de cada persona que sólo busca su propio beneficio

hacia un objetivo social óptimo, el soberano obra como un pie perfectamente visible que dirige, a puntapiés bien colocados, los esfuerzos de cada individuo al mismo objetivo social. Cada artificio realiza la misma tarea y asegura la coincidencia de un equilibrio en el cual cada persona maximiza su utilidad esperada atendiendo a las acciones de sus semejantes, con un objetivo óptimo en el cual cada persona obtiene la utilidad máxima compatible con las utilidades de sus semejantes. Cada artificio afecta las condiciones en las cuales se desarrolla la interacción y deja que cada individuo maximice libremente su utilidad atendiendo a esas condiciones. Por supuesto, el soberano aparece como una restricción a la libertad de cada persona mientras que el mercado no lo hace, pero esa es la diferencia entre la visibilidad y la invisibilidad; el soberano establece visiblemente las condiciones que concilian el interés de cada persona con el interés de sus pares, mientras que el mercado establece esas condiciones sencillamente en virtud de su estructura.

El soberano hace que la moral, entendida como una restricción al empeño que cada persona pone para maximizar su utilidad, sea tan innecesaria como en el mercado. Los problemas que le crean las externalidades a los maximizadores de la utilidad motivaron nuestra indagación moral. Adam Smith nos recuerda en qué condiciones están ausentes las externalidades, de modo tal que el mercado asegura que cuando cada persona es libre, la conducta maximizadora conduce a un resultado óptimo. Thomas Hobbes introduce la figura del soberano, quien restringe las opciones de cada individuo aun cuando están presentes las externalidades. Podemos conservar la idea de la justicia como aquello que expresa la exigencia de imparcialidad que tienen que tener los principios que regulan la interacción social, pero que ya no expresa una restricción a la maximización individual. Así parecería que la economía y la política resuelven entre ellas nuestro problema sin necesidad de apelar a la moral.

Pero el soberano de Hobbes no apela al mercado y tiene una buena razón para no hacerlo. La mano invisible es una solución sin costo alguno a los problemas de la interacción natural, pero el pie visible es una solución muy costosa. Aquellas personas sujetas al soberano de Hobbes no pueden en realidad alcanzar un resultado óptimo; cada uno paga una porción de los costos

necesarios para obligar a todos a adherirse a los acuerdos y esos costos hacen que el resultado sea subóptimo. Aun cuando supongamos que el poder no corrompe y que el soberano sea el instrumento perfecto con que cuentan sus súbditos que actúan sólo guiados por su propio interés, cada individuo podría esperar un resultado mejor si todos se adhirieran voluntariamente a los acuerdos, de modo tal que la observancia forzosa y sus costos fueran innecesarios. Pagamos un alto precio si somos en realidad criaturas que no aceptamos racionalmente ninguna restricción interna a la búsqueda de nuestra propia utilidad y que consecuentemente somos capaces de escapar al estado de la naturaleza en aquellas circunstancias en las que las externalidades están irremediablemente presentes, sólo mediante artificios políticos y no morales. Pero, si pudiéramos acatar voluntariamente nuestros acuerdos acordados racionalmente, podríamos ahorrarnos ese precio.

No suponemos que el acatamiento voluntario pueda eliminar la necesidad de que existan instituciones y prácticas sociales con sus consecuentes costos. Pero ese consentimiento podría eliminar la necesidad de que existan algunas de esas instituciones creadas para forzar la observancia. No es posible eliminar la toma perentoria de decisiones, pero nuestro ideal sería una sociedad en la cual no fuera necesario forzar coercitivamente tales decisiones. De manera más realista, suponemos que tal observancia forzosa es necesaria para crear y mantener aquellas condiciones en las cuales los individuos pueden racionalmente esperar de sus semejantes el grado de acatamiento necesario para considerar deseable el propio acatamiento voluntario. Las restricciones morales internas obran con el fin de asegurar el acatamiento en condiciones de seguridad establecidas por las restricciones políticas externas. Pero antes de pretender que el lector acepte este punto de vista debemos mostrar lo que el Necio niega, o sea que es racional disponerse a cooperar y por lo tanto aceptar las restricciones morales internas. El argumento que presenta Hobbes de que aquellos que no estén dispuestos a hacerlo no pueden racionalmente ser recibidos dentro de la sociedad es la base sobre la que desarrollaremos nuestro análisis.

2.1 El Necio, y aquellos que comparten su concepción de la

razón práctica, deben suponer que existen potencialidades para la cooperación con las que cada persona debería estar racionalmente de acuerdo si espera que el pacto se lleve realmente a cabo, pero que tales potencialidades permanecen en el plano hipotético puesto que cada individuo espera racionalmente que cada participante, quizás él mismo, quizás algún otro, decida no adherirse al acuerdo. En el capítulo 5 sostuvimos que la cooperación es racional si cada cooperador puede esperar una utilidad aproximadamente igual a aquella que se le asignaría por el principio de concesión relativa minimáxima. El Necio no discute la necesidad de esa condición, pero niega que sea suficiente. E insiste en afirmar que para que sea racional acatar un acuerdo de cooperación, el individuo debe poder esperar de la cooperación una utilidad que no debe ser menor de la que podría obtener si violara el acuerdo. Y luego sostiene que para que sea racional que un individuo acepte participar de la cooperación (aunque ese individuo no necesariamente debe considerar racional su propio acatamiento) éste debe creer que es racional para los demás dar su consentimiento. Suponiendo que todos fueran racionales, estuvieran plenamente informados y tuvieran las expectativas correctas, el Necio supone que la cooperación sólo llega a realizarse si cada persona espera de ella una utilidad no menor que la que le proporcionaría el no acatamiento del acuerdo cooperativo. Los beneficios que podrían alcanzarse mediante acuerdos cooperativos que no le proporcionen a cada individuo por lo menos la utilidad que le brinda el no acatamiento permanecerán para siempre fuera del alcance de los seres humanos racionales, es decir nos serán siempre negados porque nuestra racionalidad misma nos haría violar los acuerdos necesarios para alcanzar esos beneficios. De modo que tales acuerdos nunca han de llevarse a cabo.

El Necio rechaza el punto de vista aparentemente corriente de que, cuando no se presentan circunstancias imprevistas ni hay una representación errada de los términos, es racional acatar un acuerdo si es racional establecerlo. Y afirma que quienes sostienen esta idea pasan por alto las implicaciones de la concepción maximizadora de la racionalidad práctica. Al elegir, un individuo adopta una posición en el presente y vislumbra la utilidad esperada que ha de resultar de cada acción

posible. Lo que ha ocurrido puede afectar esa utilidad; el hecho de que uno haya establecido un acuerdo puede afectar la utilidad que uno espera de hacer o no hacer aquello que mantendría el acuerdo. Pero lo que ocurrió no proporciona por sí mismo una razón para la elección. El hecho de que un individuo haya tenido razones para establecer un acuerdo, sólo puede darle razones para mantenerlo afectando la utilidad del acatamiento. Pensar lo contrario equivale a rechazar la maximización de la utilidad.

Comencemos nuestra respuesta al Necio recordando la distinción presentada en 5.1.3 entre una estrategia individual y una estrategia conjunta.[18] Una estrategia individual es una lotería que obra sobre las posibles acciones de un actor único. Una estrategia conjunta es una lotería que obra sobre los posibles resultados. Los cooperadores pueden elegir estrategias conjuntas.

Podemos concebir la participación en una actividad cooperativa (tal como una cacería, en la cual cada cazador desempeña una función particular coordinada con las de los demás) como la implementación de una estrategia conjunta única. Y podemos extender este concepto e incluir la participación en una práctica, tal como hacer y mantener promesas, en la cual la conducta de cada persona se basa en el acatamiento de los demás de esa práctica.

Un individuo no puede estar seguro de que actúa dentro de una estrategia conjunta porque depende no sólo de sus propias intenciones sino también de las de aquellos con quienes participa de la interacción. Pero podemos decir que un individuo funda sus actos en una estrategia conjunta en la medida en que elija intencionalmente lo que la estrategia exige de él. Normalmente, por supuesto, una persona basa sus actos en una estrategia conjunta sólo si espera que aquellos con quienes participa de la interacción hagan lo propio, de modo que en realidad esa persona espera estar actuando dentro de una estrategia conjunta. Pero no necesariamente tenemos que trasladar tal expectativa a la concepción de basar los actos de una persona en una estrategia conjunta.

Una persona coopera con sus semejantes únicamente si basa sus actos en una estrategia conjunta; aceptar cooperar es

aceptar emplear una estrategia conjunta antes que una estrategia individual. El Necio afirma que es racional cooperar sólo si la utilidad que ese individuo espera de actuar según una estrategia cooperativa es por lo menos igual a la utilidad que podría esperar si en cambio obrara según la mejor estrategia individual. Esto da por tierra con el objetivo de la cooperación que es en efecto sustituir las estrategias individuales por una estrategia conjunta en aquellas situaciones en las que esa sustitución resulta beneficiosa para todos.

Una estrategia conjunta es plenamente racional sólo si proporciona un resultado óptimo o, en otras palabras, sólo si le proporciona a cada persona que actúa siguiendo esa estrategia, la máxima utilidad compatible en esa situación con la utilidad que recibe cada una de las demás personas que actúan siguiendo la misma estrategia. Por consiguiente, podemos decir que una persona que actúa siguiendo una estrategia conjunta racional maximiza su utilidad, sujeta a la restricción establecida por las utilidades que esa estrategia les brinda a cada uno de los demás. Una estrategia individual es racional si y sólo si maximiza la utilidad de esa persona, atendiendo a las *estrategias* adoptadas por las demás personas; una estrategia conjunta es racional sólo si (pero no si y sólo si) maximiza la utilidad de un individuo atendiendo a las *utilidades* que reciben las demás personas.

Digamos que un maximizador *directo* es una persona que busca maximizar su utilidad atendiendo a las estrategias de aquellos con quienes participa en la interacción. Por otro lado, un maximizador *restringido* es una persona que en ciertas situaciones busca maximizar su utilidad atendiendo no a las estrategias sino a las utilidades de los demás que participan de la interacción. El Necio acepta la racionalidad de la maximización directa. Nosotros, defendiendo la condición *A'* de la racionalidad estratégica (establecida en 1.1) aceptamos la racionalidad de la maximización restringida.

Un maximizador restringido tiene una disposición condicional a basar sus acciones en una estrategia conjunta sin considerar si alguna estrategia individual podría proporcionarle una utilidad esperada mayor. Pero no toda restricción puede ser racional; debemos especificar las características de la disposición condicional. Por consiguiente, un maximizador restringi-

do como: (*i*) alguien que está condicionalmente dispuesto a basar sus acciones en una estrategia o práctica conjunta, si cree que los demás basan también sus actos en esa estrategia y espera una utilidad no menor que la que podría proporcionarle el hecho de que todos empleen estrategias individuales y aproximadas a la que podría esperar del resultado cooperativo determinado por las concesión relativa minimáxima; (*ii*) alguien que realmente obra con esa disposición condicional, si espera que la utilidad sea mayor que la que podría esperar en el caso de que todos emplearan estrategias individuales. O, en otras palabras, un maximizador restringido está dispuesto a cooperar de formas que, si cuentan con la adhesión de todos, han de proporcionar resultados que el maximizador considere beneficiosos y no injustos, y si coopera debe esperar que una práctica o actividad real sean beneficiosas. Al determinar esto último el maximizador restringido debe tener en cuenta la posibilidad de que algunas personas no puedan o no quieran obrar cooperativamente. De aquí en adelante, salvo que establezcamos específicamente lo contrario, al hablar de un maximizador restringido estaremos hablando de un actor con esa disposición particular.

Es necesario destacar tres puntos de nuestra caracterización de la maximización restringida. El primero es que un maximizador restringido está condicionalmente dispuesto a obrar no solamente siguiendo la única estrategia conjunta que podría prescribir una negociación racional, sino también cualquier estrategia conjunta que le proporcione una utilidad semejante a la que podría esperar de la plena cooperación racional. La cantidad de estrategias conjuntas aceptables no queda especificada y así debe ser. La idea es que en la interacción real es razonable aceptar arreglos cooperativos que se aproximen al ideal de racionalidad y justicia plena, siempre que se aproximen lo suficiente. Llegada a cierto punto, una persona, por supuesto, decide ignorar una estrategia conjunta, aun cuando ajustarse a ella pudiera proporcionarle una utilidad mayor de la que cabría esperar si cada uno empleara una estrategia individual, porque esa persona espera así obtener la adhesión a otra estrategia conjunta que al ser más justa es también más favorable para sí. En qué preciso momento la persona decide hacer esto es algo que no trateremos de establecer aquí. Sencillamente defendemos

una concepción de la maximización restringida que no exige que todas las estrategias conjuntas sean ideales.

Por consiguiente, la maximización restringida vincula la idea de una moral basada en el acuerdo con la práctica moral real. Suponemos que puede entenderse que los principios morales representan estrategias conjuntas prescriptas a cada persona como parte de acuerdos cooperativos en desarrollo que constituyen la sociedad. Esos principios le exigen a cada persona que refrene la búsqueda directa de su utilidad máxima a fin de alcanzar resultados mutuamente ventajosos y razonablemente justos. En general, los principios morales reales no son aquellos con los que deberíamos estar de acuerdo en una negociación plenamente racional, pero es razonable adherirse a ellos en la medida en que se aproximen razonablemente a los principios ideales. Podemos defender los principios morales reales tomando como referencia los acuerdos cooperativos ideales y cuanto más se aproximen esos principios, más fuerza tendrá la defensa. Por supuesto, no suponemos que nuestros principios morales reales deriven históricamente de una negociación, pero en la medida en que las restricciones que imponen le resulten aceptables a un maximizador restringido, podemos incluirlos dentro del marco de una moral racionalizada por la idea del acuerdo.

El segundo punto se refiere a que un maximizador restringido no funda sus actos en una estrategia conjunta a no ser que pueda alcanzarse un resultado aproximadamente justo y óptimo si todos lo hacen. Su disposición a cooperar es condicional, depende de la expectativa que el individuo tenga del beneficio que ha de obtener comparado con la utilidad que le cabría esperar de una situación en la que nadie cooperara. De modo que ese individuo debe estimar la probabilidad de que las demás personas implicadas en la futura práctica o interacción actúen de manera cooperativa y calcular, no la utilidad que podría esperar si todos cooperaran, sino la utilidad que podría esperar si él cooperara atendiendo a la estimación que hizo de la probabilidad de que los demás también cooperen. Sólo si esta utilidad es mayor que la que cabría esperar de la no cooperación universal, la disposición condicional a restringir la búsqueda del propio interés se manifestaría realmente en la decisión de basar las acciones en una estrategia cooperativa conjunta.

De modo que, al interactuar con personas a las que considera maximizadores directos, el maximizador restringido no se entrega en manos de los demás basando sus acciones en una estrategia conjunta que él desearía que todos acepten, sino que, a fin de evitar que lo exploten, se conduce como un maximizador directo y actúa siguiendo una estrategia individual que maximiza su propia utilidad atendiendo a las estrategias que espera que empleen los demás. Un maximizador restringido considera razonablemente cierto que se encuentra entre personas con una disposición semejante a la suya, antes de restringir realmente la búsqueda directa de su máxima utilidad.*

Pero téngase en cuenta que un maximizador restringido puede verse obligado a actuar de un modo tal que le hubiera convenido no participar de la cooperación. Puede encontrarse comprometido en una actividad cooperativa que, dada la voluntad de los demás de cumplir su parte, parece justa y beneficiosa, pero que para ser realmente así le exige actuar de un modo que le provoca una pérdida mayor de lo que nunca estuvo dispuesto a arriesgar. Esta persona debería estar aun dispuesta a acatar el acuerdo y a obrar de un modo que le provocaría una verdadera desventaja, porque, atendiendo a sus creencias *ex ante* referentes a la disposición de sus pares y a las perspectivas de beneficio, la participación en esa actividad debía proporcionarle una utilidad esperada mayor que la no participación.

Y esto nos lleva al tercer punto: la maximización restringida no es la maximización directa en su apariencia más efectiva. El maximizador restringido no es meramente la persona que, pudiendo ver más allá que sus semejantes, sirve a su interés general sacrificando los beneficios inmediatos de ignorar las estrategias conjuntas y violar los acuerdos cooperativos a fin de obtener beneficios de largo plazo gracias a la confianza depositada en él por los demás.[19] Una persona semejante no acepta ninguna restricción verdadera. El maximizador restringido no razona de manera más efectiva sobre cómo maximizar su utilidad, sino que razona de un modo diferente. Podemos comprenderlo más claramente si consideramos cómo afronta cada uno la decisión de basar o no sus acciones en una estrategia conjunta. El maximizador restringido considera (*i*) si el resultado (en el caso de que todos siguieran la estrategia conjunta) ha

de ser aproximadamente justo y óptimo y (*ii*) si el resultado que puede realmente esperar (si sigue la estrategia conjunta) le proporcionaría una utilidad mayor que la que le brindaría la no cooperación universal. Si ambas condiciones se cumplen, el maximizador restringido obra siguiendo la estrategia conjunta. El maximizador directo considera simplemente si el resultado que realmente puede esperar (en el caso de que siga la estrategia conjunta) ha de proporcionarle una utilidad mayor que el resultado que podría esperar si decidiera obrar siguiendo cualquier estrategia alternativa; considerando por supuesto tanto los efectos de largo como de corto alcance. Sólo si se cumple esta condición el maximizador directo basa sus actos en una estrategia conjunta.

Consideremos una interacción puramente aislada, en la cual ambas partes saben que lo que decidan no ha de influir en lo que luego le ocurra a cada uno en otras interacciones. Supongamos que la situación tiene la estructura ya familiar del dilema del prisionero; para cada actor es más beneficiosa la cooperación mutua que la no cooperación, pero cada uno se beneficia más con la no cooperación independientemente de lo que haga el otro. En una situación semejante, el maximizador directo decide no cooperar. Un maximizador restringido en cambio decide cooperar si, teniendo en cuenta su apreciación respecto de lo que ha de decidir el otro, su utilidad esperada es mayor que la utilidad que podría esperar de un resultado no cooperativo.

Por lo tanto, los maximizadores restringidos pueden obtener beneficios cooperativos inaccesibles a los maximizadores directos, por largas que sean las miras de estos últimos. Pero los maximizadores directos pueden, en ocasiones, explotar a los maximizadores restringidos incautos. Ambos suponen que su disposición es racional. Pero, ¿cuál está en lo cierto?

2.2 A fin de demostrar la racionalidad de la maximización convenientemente restringida resolveremos un problema de elección racional. Consideramos lo que un individuo racional elija — ya sea adoptar una maximización directa, ya sea adoptar una maximización restringida— como su disposición para la conducta estratégica. Aunque esa elección se refiera a la interacción, hacerla no significa comprometerse en la interac-

ción. Si considera que las disposiciones de los demás son algo fijo, ese individuo razona tomando como parámetro su propia disposición más ventajosa. Por consiguiente, compara la utilidad esperada que le depara disponerse a maximizar la utilidad (dadas las elecciones estratégicas esperadas de los demás) con la utilidad que le depara disponerse a cooperar con los demás a fin de obtener resultados aproximadamente justos y óptimos.

Para poder elegir entre esas disposiciones, una persona necesita considerar únicamente aquellas situaciones en las cuales tales disposiciones han de provocar diferentes comportamientos. Si ambas disposiciones pudieran expresarse en una estrategia individual maximizadora, o si ambas llevaran a la persona a basar sus actos en la estrategia conjunta que esa persona espera de los demás, las expectativas de utilidad de las dos disposiciones serían idénticas. Pero si la disposición restrictiva se expresara en basar los actos en una estrategia conjunta en tanto que la disposición a maximizar directamente se expresara en el hecho de abandonar la estrategia conjunta, las expectativas de utilidad diferirían. Sólo deben considerarse las situaciones que dan lugar a tales diferencias. Y esas situaciones tienen que cumplir dos condiciones. Primero, deben proporcionar la perspectiva de una cooperación mutuamente beneficiosa y justa, pues de otro modo la restricción sería inútil. Y en segundo lugar, deben brindar cierta perspectiva de que el hecho de abandonar la estrategia conjunta sea individualmente beneficioso, pues de lo contrario, no sería necesaria ninguna restricción para alcanzar los beneficios mutuos. Suponemos, pues, que tenemos un individuo que está considerando qué disposición adoptar en situaciones en las cuales su utilidad esperada es u si cada persona obra según una estrategia individual, que es u' si todos obran según una estrategia conjunta cooperativa, y que es u'' si él obra según una estrategia individual y los demás basan sus actos en una estrategia conjunta cooperativa; suponemos también que u es menor que u' (de modo tal que nuestro individuo se beneficia de la cooperación como lo exige la primera condición) y que a su vez u' es menor que u'' (de modo que el individuo se beneficia abandonando la estrategia conjunta como lo exige la segunda condición).

Consideremos ahora estos dos argumentos que el individuo puede esgrimir:

Argumento (1): Supongamos que adopto la maximización directa. En ese caso espero que los demás basen sus actos en una estrategia conjunta, abandono esa estrategia y adopto mi mejor estrategia individual, con lo cual espero una utilidad u''. Si espero que los demás obren según estrategias individuales, yo hago lo mismo y espero una utilidad u. Si la probabilidad de que los demás basen sus acciones en una estrategia conjunta es p, luego mi utilidad esperada general es $[pu'' + (1 - p)u]$.
Supongamos que adopto la maximización restringida. En ese caso, si espero que los demás basen sus actos en una estrategia conjunta, yo hago lo mismo y espero una utilidad u'. Si espero que los demás sigan estrategias individuales, también hago yo lo mismo y espero una utilidad u. Por consiguiente mi utilidad esperada general es $[pu' + (1 - p)u]$.
Puesto que u'' es mayor que u' $[pu'' + (1 - p)u]$ es mayor que $[pu' + (1 - p)u]$, para cualquier valor de p diferente de 0 (y si $p=0$, los dos resultados son iguales). Por consiguiente, a fin de maximizar mi expectativa general de utilidad adopto la maximización directa.

Argumento (2): Supongamos que adopto la maximización directa. En ese caso debo esperar que los demás empleen estrategias individuales maximizadoras al participar conmigo de la interacción; de modo que yo hago lo mismo y espero una utilidad u.
Supongamos que adopto la maximización restringida. En ese caso, si los demás están condicionalmente dispuestos a la maximización restringida, puedo esperar que basen sus actos en una estrategia cooperativa conjunta al estar en interacción conmigo; por lo tanto yo hago lo mismo y espero una utilidad u'. Si ellos no tienen esa disposición, yo empleo una estrategia maximizadora y espero una utilidad u como antes. Si la probabilidad de que los demás estén dispuestos a adoptar la maximización restringida es p, mi utilidad esperada general es pues $[pu' + (1 - p)u]$.
Puesto que u' es mayor que u $[pu' + (1 - p)u]$ es mayor que u para cualquier valor que pueda tener p distinto de 0 (y si $p=0$, las dos utilidades son iguales). Por consiguiente, a fin de maximizar mi expectativa de utilidad general, yo debería optar por la maximización restringida.

Como estos dos argumentos llegan a conclusiones opuestas, no pueden ser ambos correctos. El primero tiene la forma de un argumento dominante. En cualquier situación en la que los demás obren de manera no cooperativa, el individuo puede

esperar la misma utilidad si está dispuesto a optar por la maximización directa que si está dispuesto a optar por la maximización restringida. En cualquier situación en la que los demás obren de manera cooperativa, el individuo puede esperar una mayor utilidad si está dispuesto a la maximización directa. Por lo tanto debería optar por la maximización directa. Pero este argumento sólo podría ser válido si la probabilidad de que los demás obren de manera cooperativa, como el mismo argumento supone, fuera independiente de la propia disposición del individuo. Y ése no es el caso. Puesto que las personas dispuestas a la cooperación, sólo obran de manera cooperativa con aquellas a las que suponen igualmente dispuestas, un maximizador directo no tiene las oportunidades de beneficiarse que se les presentan a los maximizadores restringidos. Por consiguiente, el argumento (1) es incorrecto.

El argumento (2) tiene en cuenta lo que el argumento (1) ignora: la diferencia entre la manera en que interactúan los maximizadores restringidos con aquéllos similarmente dispuestos y la manera en que interactúan con los maximizadores directos. Sólo son racionalmente aceptables como partes de los acuerdos aquellos individuos dispuestos a cumplirlos. Los maximizadores restringidos pueden establecer acuerdos beneficiosos con sus pares que los maximizadores directos no pueden establecer, no porque no estén dispuestos a acordar, sino porque no serían admitidos como partes del acuerdo debido a la disposición que tienen a violarlo. Los maximizadores directos están dispuestos a sacar ventaja de sus semejantes cada vez que tengan la oportunidad de hacerlo. Con las mismas oportunidades los maximizadores directos deberían necesariamente obtener mayores beneficios. Un argumento dominante establece esto. Pero puesto que difieren en sus disposiciones, los maximizadores directos y los maximizadores restringidos difieren también en las oportunidades y los últimos salen beneficiados.

Pero desafortunadamente el argumento (2) contiene un supuesto que no está sustentando. Las expectativas que una persona tiene respecto de cómo han de interactuar los demás con él dependen estrictamente de la disposición que él elija, sólo si los demás conocen esa elección. Lo que hemos mostrado es que,

si el maximizador directo y el maximizador restringido aparecen con sus verdaderos rostros, el maximizador restringido ha de salir beneficiado. Pero, ¿necesariamente han de mostrarse así? El Necio puede estar de acuerdo con que (ante la presión de nuestro argumento y su paralelo en el segundo argumento que le atribuimos a Hobbes) lo que debemos preguntarnos es no si es racional o no mantener los pactos (particulares), sino si es o no racional estar (generalmente) dispuesto a cumplir los pactos. Y el Necio ha de reconocer que no puede ganar defendiendo directamente la causa del maximizador directo. Pero, ¿no podría acaso ganar el Necio si suma a la maximización directa la apariencia de la restricción? ¿Acaso no es el argumento último del Necio que la persona verdaderamente prudente, el maximizador de la utilidad plenamente racional debe tratar de parecer un confiable cumplidor de sus acuerdos? Porque de ese modo no quedaría excluido de los acuerdos cooperativos de sus pares, sino que sería bien recibido como un socio, mientras que en realidad él espera las oportunidades de beneficiarse a expensas de los demás, y preferiblemente sin que los demás lo sepan, de modo tal de poder conservar su apariencia de restricción y confiabilidad.

Existe un modo fácil de desbaratar esa maniobra. Puesto que nuestro argumento está destinado a aplicarse a personas idealmente racionales, podemos añadir sencillamente otro supuesto ideal y considerar que nuestras personas son *transparentes*.[20] Cada uno de los actores tiene directamente conciencia de la disposiciones de sus semejantes y por lo tanto, sabe si quienes participan con él en la interacción son maximizadores directos o maximizadores restringidos. El engaño es imposible. El Necio debe aparecer como realmente es.

Pero suponer la transparencia parece quitarle a nuestro argumento gran parte de su interés. Queremos relacionar nuestros supuestos ideales con el mundo real. Si la maximización restringida sólo vence a la maximización directa en el caso de que todas las personas sean transparentes, fracasaríamos al tratar de mostrar que en condiciones posibles, realistas y verdaderas las restricciones morales son racionales. Refutaríamos la posición del Necio pero el precio que tendríamos que pagar sería quitarle a nuestra refutación toda aplicación práctica.

Con todo, la transparencia demuestra ser un supuesto más fuerte de lo que exige nuestro argumento. Podemos apelar en cambio a un carácter *traslúcido* más realista y suponer que las personas no son ni transparentes ni opacas, de modo tal que la disposición que tengan a la cooperación o a la no cooperación pueda ser atribuida por los demás, no con certeza, pero sí con más peso que el de la mera conjetura. Los seres opacos estarían condenados a buscar soluciones políticas a aquellos problemas de la interacción natural que no pueda solucionar el mercado. Pero mostraremos que las personas traslúcidas (todo lo traslúcidas que podamos razonablemente considerarlas) disponen racionalmente de soluciones morales.

2.3 Si las personas fueran traslúcidas, los maximizadores restringidos (MR) a veces no podrían reconocerse recíprocamente, de modo tal que participarían de manera no cooperativa en la interacción aun cuando la cooperación podría haber sido mutuamente beneficiosa. Los MR a veces no podrían identificar a los maximizadores directos (MD), de modo tal que obrarían con ellos de manera cooperativa; si los MD identificaran correctamente a los MR podrían sacar ventaja de ellos. Los MR traslúcidos deben esperar salir menos beneficiados de la interacción que los MR trasparentes; los MD traslúcidos deben esperar salir más beneficiados que los MD transparentes. Aunque sería racional que una persona elija ser un MR, si fuera transparente; ya no lo es necesariamente si esa persona es sólo traslúcida. Examinemos las condiciones en las cuales disponerse a la maximización restringida es racional para las personas traslúcidas y preguntémonos si ésas son (o si pueden ser) las condiciones en las cuales nos encontramos.

Lo mismo que en la subsección anterior, tenemos que considerar únicamente aquellas situaciones en las cuales los MR y los MD pueden obrar de manera diferente. Son situaciones que brindan tanto la perspectiva de una cooperación mutuamente beneficiosa (comparada con la no cooperación), como la perspectiva de una renuncia a la cooperación individualmente beneficiosa (comparada con la cooperación). Simplifiquemos esto y supongamos que se llega al resultado no cooperativo a menos que (*i*) aquellos que participan en la interacción sean MR que logran el reconocimiento mutuo (con lo cual se llegaría al

resultado cooperativo), o que (*ii*) entre aquellos que participan de la interacción haya MR que no reconozcan a los MD pero que sí sean reconocidos por éstos (con lo cual el resultado les proporcionaría a los MD los beneficios del abandono individual de la cooperación y cargaría sobre los MR los costos de haber basado erróneamente sus actos en una estrategia cooperativa y haber permitido que los MD sacaran ventaja). Ignoramos aquí la ventaja inadvertida que pueda sacar un MR al confundir a los otros MR con MD.

De modo que hay cuatro posibilidades: la no cooperación, la cooperación, la defección de la cooperación y la explotación (así podemos llamar al resultado que obtienen las personas que, siendo consideradas partes, abandonan la estrategia conjunta en las que basan sus actos). En la situación típica asignamos un valor 1 al abandono, un valor u'' (menor que 1) a la cooperación, un valor u' (menor que u'') a la no cooperación y un valor 0 (menor que u') a la explotación. Y ahora presentamos las tres probabilidades. La primera, p, es la probabilidad de que los MR logren el reconocimiento mutuo y que por lo tanto cooperen con éxito. La segunda, q, es la probabilidad de que los MR no reconozcan a los MD, pero sean reconocidos por éstos, de modo tal que lo que ha de resultar es o bien la defección o bien la explotación. La tercera, r, es la probabilidad de que un miembro de la población seleccionado al azar sea un MR. (Suponemos que todos son o bien MR o bien MD, de modo tal que la probabilidad de que una persona seleccionada al azar sea un MD es $(1 - r)$.) Los valores de p, q y r deben estar, por supuesto entre 0 y 1.

Calculemos ahora las utilidades que les cabría esperar a los MR y a los MD en las situaciones que les proporcionan tanto la perspectiva de la cooperación mutuamente beneficiosa como la perspectiva del abandono de la cooperación individualmente beneficiosa. Un MR ha de esperar la utilidad u' a menos que (*i*) tenga la suerte de cooperar con otro MR o que (*ii*) le toque ser explotado por MD. La probabilidad de (*i*) es la probabilidad combinada de que esa persona participe de la interacción con un MR, r, y de que ambos logren reconocerse mutuamente, p, o rp. En este caso nuestro individuo gana $(u'' - u')$ sobre su expectativa no cooperativa u'. Por lo tanto el efecto de (*i*) es aumentar su expectativa de utilidad mediante un valor $[rp\ (u'' - u')]$. La

probabilidad de (*ii*) es la probabilidad combinada de que la persona participe en la interacción con un MD, $1 - r$, y que no logre reconocerlo, aunque el MD lo reconozca a él como un MR, q, o $(1 - r)\,q$. En este caso el MR recibe 0, de modo tal que pierde su expectativa no cooperativa u'. Por lo tanto el efecto de (*ii*) es reducir su expectativa de utilidad mediante un valor $[(1 - r)qu']$. Al tener en cuenta tanto (*i*) como (*ii*), un MR espera la utilidad $\{u' + [rp(u'' - u')] - (1 - r)qu'\}$.

Un MD espera la utilidad u' a menos que explote a un MR. La probabilidad de que ocurra esto es la probabilidad combinada de que esa persona participe de la interacción con un MR, r, y lo reconozca como tal mientras que éste no logra reconocerlo a él como un MD, q o rq. En este caso el MD gana $(1 - u')$ sobre su expectativa no cooperativa u'. Por lo tanto el efecto es aumentar su expectativa de utilidad mediante un valor $[rq(1 - u')]$. Un MD espera pues la utilidad $\{u' + [rq\,(1 - u')]\}$.

Unicamente es racional disponerse a la maximización restringida si y sólo si la utilidad esperada por un MR es mayor que la utilidad esperada por un MD que la obtiene si y sólo si p / q es *mayor* que $\{(1 - u')/(u'' - u') + [(1 - r)u']/[r(u'' - u')]\}$.

El primer término de esta expresión, $[(1 - u')/(u'' - u')]$, relaciona la ganancia que produce la defección de la cooperación con la ganancia que se obtiene mediante la cooperación. El valor del abandono es por supuesto mayor que el de la cooperación, de modo que este término es mayor que 1. El segundo término, $\{[(1 - r)u']/[r(u'' - u')]\}$, depende en cuanto a su valor de r. Si $r = 0$ (por ejemplo si no hay MR en la población), su valor es pues infinito. A medida que r aumenta, el valor de la expresión se reduce, hasta que si $r = 1$ (por ejemplo, si sólo hay MR en la población) su valor es 0.

Ahora podemos sacar dos importantes conclusiones. Primero es racional disponerse a la maximización restringida sólo si la razón p / q (por ejemplo la razón entre la probabilidad de que una interacción que incluya MR resulte cooperativa y la probabilidad de que una interacción de la que participen tanto MR como MD incluya casos de explotación y de abandono) es mayor que la razón entre la ganancia que produce la defección y la ganancia que se obtiene mediante la cooperación. Si toda la población estuviera compuesta por MR, podríamos pues reemplazar la expresión "sólo si" por la expresión "si y sólo si" en aquella declaración, pero en general ésta es sólo una condición necesaria para que disponerse a la maximización restringida sea una actitud racional.

Segundo, a medida que aumenta en la población la proporción de MR (de modo que también aumenta el valor de r), disminuye el valor de la razón p / q que es necesario para que disponerse a la maximización restringida sea una actitud racional. Cuantos más maximizadores restringidos haya, mayores han de ser el riesgo de abandono de la cooperación y de explotación que acepte racionalmente correr un maximizador restringido. Sin embargo, esos riesgos y particularmente el último deben ser relativamente pequeños.

Podemos ilustrar estas conclusiones introduciendo valores numéricos típicos para la cooperación y la no cooperación y considerando luego diferentes valores para r. Uno podría suponer que en general no hay razón para que la ganancia típica que produce la defección de la cooperación comparada con la de la cooperación sea mayor o menor que la ganancia típica que produce la cooperación comparada con la no cooperación, y tampoco habría razón para que esta última ganancia sea mayor o menor que la pérdida típica que provoca la no cooperación respecto de la explotación. Y así, puesto que la defección tiene un valor 1 y la explotación un valor 0, asignémosle a la cooperación el valor 2/3 y a la no cooperación el valor 1/3.

La ganancia que produce la defección de la cooperación, $(1 - u')$, es pues 2/3; la ganancia que produce la cooperación, $(u'' - u')$ es 1/3. Puesto que q / r debe ser superior a $\{(1 - u')/(u'' - u') + [(1 - r)u']/[r(u'' - u')]\}$ para que la maximización restringida sea racional, en nuestro caso típico la probabilidad p de que los MR cooperen con éxito debe ser más del doble de la probabilidad q de que los MD exploten a los MR, aun cuando existan grandes probabilidades r de que una persona seleccionada al azar sea un MR. Si tres de cuatro personas son MR, de modo tal que $r = 3/4$, p / q debe ser pues mayor que 7/3; si de dos personas una es un MR, p / q debe ser mayor que 3; si una de cuatro personas es un MR, p / q debe ser mayor que 5. En general p / q debe ser mayor que $2 + (1 - r)/r$, o $(r + 1)/r$.

Supongamos una población dividida igualmente entre

maximizadores restringidos y maximizadores directos. Si los maximizadores restringidos pueden cooperar con éxito en los dos tercios de sus encuentros y pueden evitar que los maximizadores directos los exploten en los cuatro quintos de sus encuentros, los maximizadores restringidos pueden pues esperar salir más beneficiados que los otros. Por supuesto, la distribución igual no ha de permanecer estable, pues sería racional para los maximizadores directos cambiar su posición. Esas personas son lo suficientemente *traslúcidas* para considerar que la moral es racional.

2.4 Un maximizador restringido está condicionalmente dispuesto a cooperar de maneras tales que, seguidas por todos, deberían proporcionar resultados aproximadamente óptimos, y justos y ese maximizador coopera de tal modo cuando realmente puede esperar beneficiarse. En las dos subsecciones anteriores, sostuvimos que un individuo se dispone racionalmente a cooperar de esas maneras si las personas son transparentes o si las personas son lo suficientemente traslúcidas y suficientemente propensas a obrar de modo semejante. Pero no apelamos explícitamente a la exigencia particular de que las prácticas y las actividades cooperativas sean aproximadamente óptimas y justas. Afirmamos que el resultado cooperativo brinda al individuo una utilidad mayor que la que puede brindarle la no cooperación, pero afirmar eso es mucho más débil que declarar que el resultado cooperativo se aproxima al resultado que exige la concesión relativa minimáxima.

Pero nótese que cuanto mayor es la ganancia que brinda la cooperación ($u'' - u'$), más disminuye el valor mínimo de p/q que es lo que da racionalidad a estar dispuesto a la maximización restringida. Podemos considerar que p/q es una medida del carácter traslúcido; cuanto más traslúcidos son los maximizadores restringidos, tanta más ventaja obtienen cooperando entre sí (aumentando p) y evitando que los maximizadores directos los exploten (disminuyendo q). Por consiguiente, como las prácticas y las actividades no alcanzan la optimidad, el valor esperado de la cooperación, u''_1 disminuye para aquellos que reciben menos que lo justo y aumenta el carácter traslúcido requerido para que la cooperación sea racional. Y como las prácticas y las actividades distan mucho de ser justas, disminuye el valor

esperado de cooperación para aquellos que obtienen menos de una participación justa, y aumenta el carácter traslúcido requerido para que la cooperación sea racional. Por lo tanto nuestro argumento apela implícitamente a la exigencia de que la cooperación proporcione resultados aproximadamente justos y óptimos.

Pero existe un argumento adicional para apoyar nuestra afirmación de que la disposición condicional a cooperar se limite a las prácticas y actividades que proporcionan resultados aproximadamente justos y óptimos. Y este argumento, como toda nuestra argumentación sobre la restricción, se refiere a la manera en que la disposición de un individuo afecta las características de las situaciones en las cuales ese individuo puede razonablemente esperar encontrarse. Diremos que la persona que está dispuesta a cooperar de maneras que, seguidas por todos, le brindan resultados aproximadamente óptimos y justos, es una persona que *acata estrechamente*. Y diremos que una persona que está dispuesta a cooperar de maneras que, seguidas por todos, le brindan meramente algún beneficio comparado con la no cooperación universal, es una persona que *acata ampliamente*. No vamos a negar que una persona que acata ampliamente ha de esperar beneficiarse en situaciones en las cuales los que acatan estrechamente no pueden hacerlo. Pero en muchas otras situaciones, el que acata ampliamente debe esperar que su disposición le acarree una pérdida. Porque en la medida en que los demás sepan que es alguien que acata ampliamente, tendrán razones para maximizar sus utilidades a expensas suyas, ofreciéndole "cooperar" pero dándole sólo la posibilidad de ganar apenas un poco más de lo que obtendría de la no cooperación. Puesto que una persona que acata ampliamente está dispuesta a coger cualquier beneficio que le pueda aportar una estrategia conjunta, se encuentra con muchas oportunidades, pero obtiene muy poco beneficio.

Puesto que la persona que acata estrechamente está siempre preparada a aceptar acuerdos cooperativos basados en el principio de la concesión relativa minimáxima, también está preparada para participar de manera cooperativa en cualquier interacción que pueda ser mutuamente beneficiosa fundada sobre bases igualmente racionales y justas para todos. Al

rechazar otras bases, esa persona no disminuye sus perspectivas de cooperación con otras personas racionales, y garantiza que aquellos que no estén dispuestos a participar de una cooperación justa no gocen de los beneficios de ninguna cooperación, pues hace que la injusticia les cueste cara y por lo tanto les resulte irracional.

En el próximo capítulo extenderemos el concepto del acatamiento estrecho, de modo tal que incluya no sólo la concesión relativa minimáxima, sino también la satisfacción de una norma de justicia en la posición negociadora inicial. Comprobaremos pues que en ciertas circunstancias el acatamiento estrecho establece una norma demasiado exigente. Si las instituciones de la sociedad no son ni racionales ni imparciales, la persona que acata estrechamente no puede realizar ninguna reforma efectiva de tales instituciones y además se priva de los beneficios que sin embargo esa sociedad imperfecta brinda. Por lo tanto —debemos admitirlo—, a veces la racionalidad y la imparcialidad pueden no coincidir en la decisión individual.

Pero suponemos que entre personas plenamente racionales, las instituciones, las prácticas y los acuerdos que no satisfagan las exigencias de la concesión relativa minimáxima deben manifestarse inestables. Por supuesto, puede haber algunas personas interesadas en mantener la injusticia inherente a tales estructuras. Pero entre los miembros de una sociedad que son, y saben que sus pares también lo son, racionales y que están adecuadamente informados, aquellos que consideren que reciben menos de lo que podrían esperar de una cooperación justa y óptima pueden reformar su sociedad de modo tal que ésta satisfaga las exigencias de la justicia, disponiéndose al acatamiento estrecho. Una reflexión sobre cómo se sustenta la parcialidad muestra que por más importantes que sean las medidas coercitivas, su efectividad depende finalmente del apoyo no obligado a normas que directa o indirectamente sustentan esa parcialidad, un apoyo que quizá no esté suficientemente fomentado por los inteligentes maximizadores restringidos de la utilidad individual.

2.5 Para concluir esta larga sección completaremos nuestra argumentación sobre la racionalidad de disponerse a la maximización restringida, con tres reflexiones referentes a sus

implicaciones: en lo tocante a la moral convencional, al tratamiento de los maximizadores directos y al cultivo del carácter traslúcido.

Primero, no tendríamos que suponer que el argumento defiende toda la moral convencional o todas aquellas instituciones y prácticas que tienden a alcanzar resultados justos y óptimos. Si la sociedad es, como dice Rawls "una empresa cooperativa destinada a la ventaja mutua", es pues racional pagar los costos sociales que le corresponden a cada uno, es decir, los impuestos. Pero no necesariamente es racional pagar los impuestos que nos corresponden a menos que se nos obligue efectivamente a pagarlos, en el caso en que veamos que los dólares de nuestros impuestos se utilizan (como podemos creer) para aumentar las posibilidades de una guerra nuclear y para fomentar el parasitismo corporativo e individual. Si la evasión de impuestos les parece a muchos contribuyentes una práctica racional, esto no muestra que sea irracional acatar los acuerdos justos y óptimos, sino que quizá, sólo muestra que es irracional aceptar voluntariamente la propia explotación.

En segundo lugar, no deberíamos suponer que es racional disponerse a la maximización restringida si uno no se dispone también a excluir a los maximizadores directos de los beneficios que pueden alcanzarse mediante la cooperación. Hobbes señala que aquellos que creen tener razones para violar sus pactos, no han de ser recibidos en el seno de la sociedad salvo por un error de sus pares. Si sus pares caen en tal error, pronto han de comprobar que a nadie le conviene cumplir con los pactos. Si no excluyen a los maximizadores directos de los beneficios de los acuerdos cooperativos no pueden compartir los beneficios de largo alcance que brinda la cooperación; si les permiten permanecer en el seno de la sociedad los acuerdos han de resultar ineficaces con lo cual a la larga no habrá beneficios que compartir. De modo que no habrá nada que puedan ganar los maximizadores restringidos, con lo cual estos últimos podrían unirse a los maximizadores directos en su descenso hacia la condición natural de la humanidad.

La tercera consideración se relaciona más estrechamente con los conceptos presentados en 2.3. Consideremos una vez más las probabilidades p y q, las probabilidades de que los MR

pueden reconocerse mutuamente y cooperar y la probabilidad de que los MR no puedan reconocer a los MD, pero sí puedan ser reconocidos por éstos y por lo tanto explotados. Es evidente que los MR se benefician si aumenta p y disminuye q. Y esto queda reflejado en nuestro cálculo de la utilidad esperada que les corresponde a los MR; el valor de $\{u' + [rp\ (u'' - u')] - (1 - r)\ qu'\}$ aumenta si también aumenta p o si disminuye q.

¿Qué determina los valores de p y q? La probabilidad p depende de la capacidad que tengan los MR de detectar la sinceridad de los demás MR y de revelarles su propia sinceridad. La probabilidad q depende de la capacidad que tengan los MR de detectar la falta de sinceridad de los MD y de ocultarles su propia sinceridad, y también de la habilidad que tengan los MD para detectar las sinceridad de los MR y de ocultarles su propia falta de sinceridad. Puesto que cualquier aumento de la capacidad para revelar la propia sinceridad a otros MR puede ser compensada por una disminución de la capacidad para ocultar la propia sinceridad a los MD, un MR probablemente se base ante todo en su habilidad para detectar las disposiciones de los demás antes que en su habilidad para revelar o ocultar la propia.

Un maximizador restringido racional debe tener bien desarrollada la capacidad de detectar las disposiciones de los demás. No poder desarrollar esa capacidad o descuidar su ejercicio, le impide al individuo beneficiarse con la maximización restringida. Y así la restricción puede resultar irracional. Pero lo realmente irracional es no poder cultivar o ejercitar la capacidad de detectar la sinceridad o la falta de sinceridad de los demás.

Tanto los maximizadores restringidos como los maximizadores directos deben esperar beneficiarse con su creciente habilidad para detectar las disposiciones de los demás. Pero si ambos se esfuerzan por maximizar sus capacidades (o la utilidad esperada, exenta de costos, que ha de brindarles ese perfeccionamiento) los MR pueden esperar pues mejorar su posición respecto de los MD, porque los beneficios que obtengan los MD al mejorar su capacidad de detectar a sus víctimas potenciales en definitiva han de ser compensados por las pérdidas que sufren, a medida que los MR perfeccionan su capacidad de detectar a sus explotadores potenciales. Por otro lado, aun

cuando los MR puedan no gozar de ninguna ganancia neta en sus interacciones con los MD, los beneficios que obtienen al mejorar su capacidad para detectar a los otros MR como cooperadores potenciales no quedan neutralizados por pérdidas correspondientes, sino que antes bien aumentan gracias a que los otros MR también perfeccionan a su vez la capacidad para detectarlos.

De modo que a medida que las personas mejoran racionalmente su capacidad de detectar las disposiciones de aquellos con los cuales participan en la interacción, puede esperarse que el valor de p aumente y que el valor de q permanezca relativamente constante. Pero entonces p/q aumenta, y cuanto mayor sea, menos necesidad tienen las otras circunstancias de ser favorables para que resulte racional disponerse a abrazar la maximización restringida. Aquellos que creen que la racionalidad y la moral son conceptos en conflicto pueden no advertir la importancia que tiene cultivar la capacidad de distinguir a los cooperadores sinceros de los hipócritas.

David Hume señala que si "un hombre virtuoso tuviera la desgracia de caer en una sociedad de rufianes, su visión particular de la justicia ya no le resultaría útil para su propia seguridad ni para la de los demás, de modo tal que ese hombre debería seguir sólo los dictados de la autopreservación".[21] Si nos toca vivir en una sociedad —o en un estado de naturaleza— de maximizadores directos, la maximización restringida, que es la que nos dispone a la justicia, no nos sería pues de ninguna utilidad, de modo tal que deberíamos consultar solamente los dictados directos de nuestras propias utilidades. En un mundo de Necios, no sería conveniente ser un maximizador restringido ni acatar los propios acuerdos. En tales circunstancias no sería racional ser moral.

Pero si nos hallamos en compañía de personas razonablemente justas, nosotros también tenemos razones para disponernos a abrazar la justicia. Una comunidad en la cual la mayor parte de los individuos están dispuestos a acatar prácticas y acuerdos justos y óptimos y por lo tanto a basar sus acciones en estrategias cooperativas conjuntas, ha de sustentarse por sí misma. Y un mundo semejante les ofrece a todos beneficios de los que los Necios nunca pueden gozar.

"Un razonable bribón" desafía a Hume y le dice que la afirmación "*la honestidad es la mejor estrategia,* puede ser una buena regla general, pero que tiene muchas excepciones; y el que se conduce con más sabiduría observa la regla general y al mismo tiempo saca ventaja de todas las excepciones."[22] Cándidamente Hume confiesa que "si un hombre cree que semejante razonamiento exige una respuesta, ha de ser difícil encontrar una que le parezca satisfactoria y convincente".[23] Difícil pero no imposible, agregamos nosotros. Porque la respuesta consiste en considerar la honestidad no como una estrategia, sino como una disposición. Sólo la persona verdaderamente dispuesta a la honestidad y la justicia puede esperar alcanzar plenamente sus beneficios, porque sólo este tipo de personas puede ser racionalmente admitida como parte de aquellos arreglos mutuamente beneficiosos —ya sean acuerdos reales o prácticas implícitamente acordadas— que se basan en la honestidad y la justicia, en el acatamiento voluntario. Pero una persona semejante, dada su disposición, no puede sacar ventaja de las "excepciones"; con razón considera que esa conducta es irracional. El necio y los razonables bribones, al ver los beneficios que pueden obtenerse de las excepciones, de las ventajosas brechas que pueden abrirse en la honestidad y el acatamiento, y al no ver más allá de esos beneficios, no adoptan esa disposición. Entre ellos los bribones pueden considerarse razonables, pero dicho sea entre nosotros (si no nos dejamos corromper con sus lindas palabras) ellos son sólo unos necios.

3.1 Al defender la maximización restringida, implícitamente hemos dado una nueva interpretación a la concepción maximizadora de la utilidad de la racionalidad práctica. La interpretación recibida, comúnmente aceptada por los economistas y elaborada en la teoría de la decisión bayesiana y en la teoría de los juegos de Von Neumann-Morgenstern, identifica la racionalidad con la maximización de la utilidad en el nivel de las elecciones individuales. Una elección es racional si y sólo si maximiza la utilidad esperada del actor. Nosotros identificamos la racionalidad con la maximización de la utilidad en el nivel de la disposición que cada individuo tiene al elegir. Una disposición es racional si y sólo si un individuo al adoptarla puede esperar que sus elecciones le proporcionen una utilidad no menor que la

que le proporcionarían las elecciones que podría hacer si adoptara cualquier otra disposición. Consideraremos ahora que las elecciones particulares son racionales si y sólo si expresan una disposición racional a elegir.

Podría parecer que una disposición maximizadora a elegir debería expresarse en las elecciones maximizadoras. Pero ya mostramos que no es así. El punto esencial de nuestra argumentación es que la disposición de una persona a elegir afecta la situación en la cual esa persona puede esperar encontrarse. Un maximizador directo, que está dispuesto a hacer elecciones maximizadoras, debe esperar que se lo excluya de los arreglos cooperativos que él podría considerar ventajosos. En cambio, un maximizador restringido puede esperar que se lo incluya en tales arreglos. La disposición adoptada beneficia a este último no en las elecciones que hace, sino en las oportunidades de elegir que se le ofrecen.

Hemos defendido la racionalidad de la maximización restringida como una disposición a elegir, mostrando que puede ser racionalmente elegida. Pero la idea de elegir entre varias disposiciones a elegir es un artificio heurístico que nos permite expresar la exigencia subyacente, es decir que una disposición racional a elegir sea maximizadora de la utilidad. En contextos paramétricos, la disposición a hacer elecciones directamente maximizadoras es indiscutiblemente maximizadora de la utilidad. Por consiguiente podemos utilizar el artificio de una elección paramétrica entre las disposiciones a elegir para mostrar que en los contextos estratégicos la disposición a hacer elecciones restringidas antes que elecciones directamente maximizadoras, es maximizadora de la utilidad. Sin embargo debemos poner el acento en que la clave de nuestra argumentación no es la elección misma, sino el carácter maximizador de la disposición en virtud de la cual se hace tal elección.

Pero en nuestra apelación a una elección entre las disposiciones a elegir hay una significación adicional. Porque suponemos que la capacidad para hacer tales elecciones es en sí misma una parte esencial de la racionalidad humana. Podemos imaginar seres tan mecanizados que en los contextos estratégicos sólo podrían aceptar la maximización directa como el único modo psicológicamente posible de elegir. Hobbes puede haber pensa-

do que los seres humanos estaban tan mecanizados que en realidad eran máquinas maximizadoras directas. Pero si pensó eso, seguramente se equivocó. En el corazón mismo de nuestra capacidad racional está la habilidad para enfrascarse en una reflexión autocrítica. El individuo plenamente racional es capaz de reflexionar sobre su norma de deliberación y de cambiar esa norma a la luz de la reflexión. Por lo tanto suponemos que las personas (que inicialmente pueden suponer que es racional ampliar la maximización directa de los contextos paramétricos a los contextos estratégicos) pueden reflexionar sobre las implicaciones que tiene semejante ampliación y rechazarla en favor de la maximización restringida. Tales personas estarían haciendo la verdadera elección de una disposición a elegir que hemos estado analizando en este capítulo.

Y al hacer esa elección, esas personas estarían expresando su naturaleza no sólo de seres racionales, sino también de seres morales. Si la disposición a hacer elecciones directamente maximizadoras estuviera incorporada en nosotros no podríamos restringir nuestras acciones del modo en que lo exige la moral. Con razón los filósofos morales se han negado a aceptar la interpretación recibida de la relación entre la racionalidad práctica y la maximización de la utilidad, porque reconocieron que tal interpretación no deja lugar para una restricción racional a la conducta directamente maximizadora de la utilidad, y por lo tanto tampoco deja lugar para la moral como se la entiende corrientemente. Pero luego se volvieron hacia un enfoque neokantiano de la racionalidad que los llevó a dejar de lado la idea de que aquellas consideraciones que constituyen las razones que una persona tiene para obrar deben tener cierta relación particular con dicha persona.[24] Olvidaron relacionar nuestra naturaleza de seres morales con nuestro interés cotidiano por satisfacer nuestras preferencias individuales. Pero ya mostramos cómo la moral surge de ese interés. Cuando comprendemos correctamente cómo se identifica la maximización de la utilidad con la racionalidad práctica, comprobamos que la moral es una parte esencial de la maximización.

3.2 Un detractor afirmaría que puede ser racional disponerse a la maximización restringida, pero negaría que las elecciones que entonces esa persona está dispuesta a hacer sean

racionales.[25] El detractor declara que hemos meramente desarrollado otra instancia de la racionalidad de no conducirse racionalmente. Y antes de que podamos acusarlo de haber caído en una paradoja el detractor nos presenta otras instancias adicionales.

Consideremos, dice, los costos que implica tomar una decisión. La maximización puede ser el procedimiento más confiable, pero no necesariamente es el más efectivo en lo que a los costos se refiere. En muchas circunstancias, la persona racional no pretende maximizar sino alcanzar un umbral preestablecido de satisfacción y elige el primer curso de acción de aquellos que se le presentan al espíritu que espera alcance ese umbral. En realidad, puede señalar nuestro objetor, los seres humanos, como otros animales superiores, son satisfactores naturales. Lo que nos diferencia de los demás animales es que los seres humanos tenemos una mayor flexibilidad, de modo tal que podemos elegir de maneras diferentes, pero los costos son tan altos que generalmente no es ventajoso ejercer nuestra opción, aunque sepamos que la mayor parte de nuestras elecciones no son maximizadoras.

Consideremos también la tendencia a creer en lo que se desea. Si nos ponemos a calcular el curso mejor o maximizador de acción, probablemente confundamos las verdaderas expectativas con las esperanzas. Sabiendo esto, nos protegemos y elegimos basándonos en principios fijos y nos adherimos a tales principios aun cuando creamos que sería mejor ignorarlos, porque sabemos que en esas cuestiones las apariencias a veces engañan. En realidad, puede afirmar nuestro detractor, en gran parte la moral debe entenderse no como un conjunto de restricciones a la maximización destinado a asegurar el justo beneficio mutuo, sino como un conjunto de restricciones a una conducta basada en creer lo que se desea a fin de asegurar una mayor aproximación a la maximización.

Consideremos además, dice el objetor, los beneficios que puede brindar una conducta amenazadora. Puedo inducir a alguien a realizar una acción ventajosa para mí si logro convencerlo de que si no lo hiciera, yo realizaría una acción verdaderamente costosa para él, aunque ésa no sea mi elección maximizadora de la utilidad. Un grupo de terroristas secuestra

un avión y amenazan con matar a todos los que viajan a bordo, incluso a ellos mismos, si el piloto no desvía el avión a La Habana. Las naciones amenazan con la represalia nuclear en caso de que sus enemigos las ataquen. Aunque cumplir con una amenaza sea costoso, si la amenaza surte efecto, no es necesario cargar con los costos y se obtiene el beneficio que no podría lograrse de otro modo.

Pero, continúa nuestro objetor, una amenaza puede ser efectiva sólo si es creíble. Puede ocurrir que, a fin de maximizar su credibilidad y su perspectiva de ventaja, una persona deba estar dispuesta a cumplir con sus amenazas si no logra lo que exige. Y así sería racional estar dispuesto a ejercer la coacción mediante la amenaza. Pero entonces, por paridad de razonamiento con lo que dijimos antes acerca de la maximización restringida, debemos suponer que es realmente racional cumplir con nuestras amenazas. Con seguridad deberíamos en cambio suponer que, aunque es claramente irracional cumplir con una amenaza frustrada, sería sin embargo racional disponerse a aceptar sólo esta clase de irracionalidad. Y de manera similar, deberíamos suponer que, aunque es claramente irracional restringir nuestra conducta maximizadora, sin embargo puede ser racional disponerse a aceptar esa irracionalidad.

Somos inflexibles. Aceptamos que un actor que está sujeto a cierta debilidad o imperfección, pueda considerar racional disponerse a hacer elecciones que en sí mismas no son racionales. Tal disposición puede constituir la forma más efectiva de compensar la debilidad o la imperfección. Esa disposición constituye la segunda mejor racionalidad por decirlo así. Pero, aunque consideraramos racional lograr la satisfacción, no nos parecería racional realizar la acción así elegida si, sin pagar ningún costo extra, pudiéramos descubrir cuál es la acción maximizadora. Y aunque pueda parecernos racional adherirnos a los principios como un resguardo contra la satisfacción de los deseos, no sería racional hacerlo si pudiéramos saber cuál es la acción maximizadora más allá de toda duda.

Contrastemos esto con la maximización restringida. La racionalidad de disponerse a abrazar la restricción no apela a ninguna debilidad o imperfección del razonamiento del actor; en

realidad la racionalidad es más evidente para quienes razonan perfectamente y no pueden ser engañados. La disposición a la maximización restringida supera el problema de las externalidades; y se vincula con el problema fundamental que surge de la estructura de la interacción. Y el punto principal de disponerse a la restricción es adherirse a ella sabiendo que uno no está eligiendo la acción maximizadora.

Los actores imperfectos consideran racional disponerse a tomar decisiones que no sean completamente racionales. De esto no podemos sacar ninguna conclusión referente a las disposiciones y a las elecciones del actor perfecto. Si su disposición a elegir es racional, con seguridad sus elecciones han de ser pues también racionales.

Pero, ¿qué ocurre con la coacción ejercida mediante la amenaza? Aquí no podemos estar de acuerdo con nuestro objetor; un actor perfecto puede considerar racional disponerse a presionar a otro mediante una amenaza, y si esto es así, ese actor considera pues racional cumplir una amenaza frustrada. Igualmente, un actor perfecto puede considerar racional resistir una amenaza y si esto es así, ese actor considera pues también racional resistir a cualquier precio. Ya sostuvimos en algún otro lugar que la disuasión puede constituir una estrategia racional, y que por lo tanto las elecciones disuasivas no maximizadoras también son racionales.[26]

Sin embargo, en una comunidad de personas racionales debe proscribirse la conducta amenazadora. A diferencia de la cooperación, la conducta amenazadora no promueve la ventaja mutua. Una amenaza que logra tener éxito sencillamente redistribuye los beneficios en favor del amenazador; resistir con éxito una amenaza mantiene el statuo quo. Una conducta amenazadora frustrada que tiene por resultado costosos actos de coacción o de resistencia es necesariamente no óptima. Su verdadera *raison d'être* es hacer que todos estén peor que en ausencia de ella. Cualquier persona que no esté excepcionalmente situada debe pues tener una expectativa *ex ante* de que la conducta amenazadora sea desventajosa para todos. La proscripción de la conducta amenazadora debe formar parte de un acuerdo justo y óptimo entre personas racionales; una de las restricciones impuestas por la concesión relativa minimáxima

es abstenerse de hacer amenazas. Por lo tanto, nuestra argumentación muestra que la conducta amenazadora es tanto irracional como inmoral.

Los maximizadores restringidos no han de disponerse a ejercer la coacción mediante la amenaza ni a resistir las amenazas entre ellos. Pero, hay circunstancias que están más allá del límite moral, en las cuales un maximizador restringido puede considerar racional disponerse a ejercer la coacción mediante la amenaza. Si esa persona se encuentra entre maximizadores directos, y especialmente si éstos fueran tan tontos para resistirse a la amenaza, que el primero se dispusiera a ejercer la coacción mediante la amenaza sería lo mejor posible. Entonces para él sería racional cumplir las amenazas frustradas aunque éstas no sean maximizadoras de la utilidad.

Nuestro crítico no ha llevado bien su defensa. La disposición de un actor plenamente racional lleva a tomar decisiones racionales. Nuestro argumento identifica la racionalidad práctica con la maximización de la utilidad en el nivel de la disposición a elegir y tiene en cuenta las implicaciones de esa identificación al atribuirle racionalidad a las elecciones particulares.

3.3. Concluiremos este capítulo señalando un interesante paralelo de nuestra teoría de la maximización restringida: la teoría evolutiva de Robert Trivers del altruismo recíproco.[27] Hemos dicho que una población de maximizadores restringidos sería racionalmente estable; ninguno de ellos tendría razones para disponerse a abrazar la maximización directa. De manera similar, si consideraramos la maximización restringida y la maximización directa como paralelos de las tendencias genéticas al altruismo recíproco y al egoísmo, una población de altruistas recíprocos sería genéticamente estable; un egoísta mutante estaría en desventaja evolutiva. Puesto que al no poder retribuir lo que recibe quedaría excluido de las relaciones cooperativas.

Trivers sostiene que la selección natural ha de favorecer el desarrollo de la capacidad de detectar el altruismo meramente simulado. Esto, por supuesto, se corresponde con nuestra afirmación de que los maximizadores restringidos, para poder tener éxito, deben tener la capacidad de detectar a los maximizadores

directos quienes se ofrecen a cooperar pero no son sinceros. Deben evitarse las interacciones explotadoras entre los maximizadores directos y los maximizadores restringidos.

Trivers también sostiene que la selección natural ha de favorecer el desarrollo de la culpa, como un artificio destinado a motivar el cambio de actitud futura en aquellos que son incapaces de practicar la reciprocidad.[28] En nuestra argumentación no apelamos a ninguna disposición afectiva; no queremos debilitar la posición que debemos refutar, la maximización directa, suponiendo que las personas no se sienten emocionalmente dispuestas a seguirla. Pero podemos esperar que en el proceso de socialización se hagan esfuerzos destinados a desarrollar y cultivar los sentimientos de cada persona, de modo tal que si una persona se conduce como un maximizador directo ha de sentir culpa. Podemos esperar que las prácticas sociales den forma a nuestras capacidades afectivas a fin de defender la interacción cooperativa.

Si una población de altruistas recíprocos es genéticamente estable, con seguridad una población de egoístas también lo es. Como ya hemos visto, el argumento que utilizamos para afirmar que la maximización restringida es racional se basa en la proporción de maximizadores restringidos que tenga la población. Una pequeña proporción de maximizadores restringidos puede muy bien sufrir más la explotación ejercida por los maximizadores directos no detectados que por la cooperación entre ellos, a menos que su capacidad de detectar las disposiciones de los demás sea extraordinariamente efectiva. De manera similar un altruista recíproco mutante estaría en desventaja entre los egoístas. Sus intentos de cooperar serían rechazados y esa persona se arruinaría tratando de esforzarse por participar de la cooperación.

¿Debemos llegar pues a la conclusión de que tanto los grupos de altruistas recíprocos como los grupos de egoístas permanecen estables en el mundo? No necesariamente. Los beneficios de la cooperación aseguran que en cualquier conjunto dado de circunstancias, cada miembro de un grupo de altruistas recíprocos ha de salir más beneficiado que un miembro correspondiente de un grupo de egoístas. Cada altruista recíproco debería tener una ventaja reproductiva. Los grupos de altruistas

recíprocos deberían por consiguiente aumentar en relación con los grupos de egoístas en medios en los cuales ambos grupos se ponen en contacto. Los altruistas deben vencer, no en un combate directo entre ambos (aunque la posible cooperación entre los altruistas recíprocos pueda lograr la victoria) sino en el combate indirecto por la supervivencia evolutiva en un mundo de recursos limitados.

En su análisis del argumento sostenido por Trivers, Jon Elster señala dos puntos de gran importancia que nosotros podemos relacionar con nuestro propio enfoque de la maximización restringida. El primero es "El altruismo es la posición más eficiente porque *no* deriva del autointerés calculado".[29] Esto es exactamente lo que sostenemos al final de 2.1: la maximización restringida no es la maximización directa con su apariencia más efectiva. El maximizador restringido ignora genuinamente el llamado de la maximización de la utilidad al adherirse a las prácticas cooperativas que exige la concesión relativa minimáxima. No hay simulación; si la hubiera, los beneficios de la cooperación no podrían alcanzarse plenamente.

El segundo punto es que el enfoque de Trivers "no explica las instancias específicas de la conducta altruista, tales como, digamos, la tendencia a salvar a una persona que se está ahogando. Se explican los intentos de rescate como una tendencia general a realizar actos de altruismo y esa tendencia se convierte pues en el objeto de la explicación evolutiva".[30] Precisamente del mismo modo, nuestra explicación no implica darle una justificación maximizadora de la utilidad a las elecciones específicas que se adhieren a una estrategia conjunta. Antes bien explicamos esas elecciones por una disposición general a elegir de manera justa acciones maximizadoras siempre que sea posible, y así le damos una justificación maximizadora de la utilidad a esa tendencia.

Por supuesto, no tenemos la pericia necesaria para discutir si los seres humanos están genéticamente dispuestos a adoptar una conducta maximizadora de la utilidad. Pero, si los seres humanos tienen esa disposición, podemos pues llegar a la conclusión de que disponerse a la maximización restringida aumenta la aptitud genética.

Notas

1 Hobbes, *De Cive*, cap. I, párr.7; en *Man and Citizen,* pág. 115.
2 Hobbes, *Leviathan,* cap. 14, pág. 64.
3 *Ibid.*
4 *Ibid.*, cap. 13, pág. 63.
5 *Ibid*, cap. 14, pág. 64.
6 *Ibid.*
7 *Ibid*, cap. 14, págs. 64-65.
8 *Ibid.*, cap. 14, pág. 65.
9 *Ibid.*
10 *Ibid.*, cap. 21, pág. 111.
11 *Ibid.*, cap. 15, pág. 71.
12 *Ibid.*
13 *Ibid.*, cap. 15, pág. 72.
14 *Ibid.*, cap. 15, pág.73.
15 *Ibid.*
16 Hobbes, *The Questions Concerning Liberty, Necessity, and Chance,* 1656, XIV; en Sir William Molesworth (comp.), *The English Works of Thomas Hobbes*,11 vols., Londres, 1839-1845, vol. 5, págs. 193-194.
17 Hobbes, *Leviathan,* cap. 15, pág. 72.
18 Nuestra respuesta al Necio se basa en (pero invalida) nuestro análisis desarrollado en "Reason and Maximization", *Canadian Journal of Philosophy* 4, 1975, págs. 424-433.
19 Por lo tanto la maximización restringida no es una posición paralela a tales estrategias de "ojo por diente", así llamadas, defendidas en los dilemas del prisionero reiterados. Los maximizadores restringidos pueden cooperar aun cuando ni siquiera esperen que su elección afecte situaciones futuras. Por consiguiente el tratamiento que hacemos de la cooperación no apela a la reciprocidad, como necesita hacerlo Robert Axelrod; véase "The Emergence of Cooperation among Egoists", *American Political Science Review* 75, 1981, págs. 306-318.
20 Derek Parfit fue quien me señaló que el análisis desarrollado en "Reason and Maximization" daba por supuesta la transparencia. Véase el análisis que hace Parfit de la "teoría del autointerés" en *Reasons and Persons*, Oxford, 1984, especialmente las págs. 18-19. Véase también la discusión de "Reason and Maximization" en S. L. Darwall, *Impartial Reason*, Ithaca, NY, 1983, especialmente las págs. 197-198.
21 Hume, *Enquiry*, iii , i, pág. 187.
22 *Ibid*, ix, ii, págs. 282-283.
23 *Ibid.*, ix, ii, pág. 283.
24 Véase, por ejemplo, T. Nagel, *The Possibility of Altruism*, Oxford, 1970, págs. 90-124.

25 El objetor podría ser Derek Parfit; véase *Reasons and Persons*, págs 19-23. Como su libro es de aparición muy reciente no es posible analizar aquí su argumentación.

26 Véase "Deterrence, Maximization and Rationality", *Ethics* 94, 1984, págs. 474-495; también en D. MacLean (comp.), *The Security Gamble: Deterrence Dilemmas in the Nuclear Age*, Totowa, NJ, 1984, págs. 101-122.

27 R.L. Trivers, "The Evolution of Reciprocal Altruism", *Quarterly Review of Biology* 46, 1971, págs. 35-57.

28 *Ibid.*, pág. 50.

29 J. Elster, *Ulysses and the Sirens: Studies in Rationality and Irrationality*, Cambridge, 1979, pág. 145.

30 *Ibid.*, págs 145-146.

VII

La posición negociadora inicial: los derechos y la estipulación

Había una vez, hace mucho tiempo y en tierras muy lejanas, una sociedad de amos y esclavos. A diferencia de muchas otras sociedades semejantes, ésta no se basaba en falsas apelaciones ideológicas al natural señorío de los amos y a la natural esclavitud de los esclavos. Lo que distinguía a los amos de los esclavos, como bien lo sabían ambas partes, era el poder. Si hubiesen tenido la menor oportunidad, los esclavos habrían cambiado encantados sus puestos por los de sus amos, quienes por consiguiente se cuidaban mucho de no darles nunca ni la menor oportunidad.

Pero un día, mientras un grupo de amos bebían ociosamente los habituales *gin tonics* de la tarde, uno de los hombres más jóvenes, recientemente egresado de la universidad, interrumpió las observaciones triviales de los mayores y sostuvo lo siguiente:

> Señores, hemos sido unos necios durante todos estos años. Los magros placeres con que nos hemos confortado [en ese instante hizo chasquear los dedos a fin de indicar que le volvieran a llenar la copa] no han sido sino sombras de los lujos de los que podríamos haber gozado, si no hubiésemos estado malgastando nuestros recursos ejerciendo la coacción sobre nuestros esclavos. Y esos esclavos han estado realizando sus tareas de mala gana y sin poner el debido cuidado [¡Maldición, muchachito! ¿Dónde está la cascarilla de limón de mi *gin tonic*?] ante la amenaza de nuestros azotes y cadenas, en lugar de servirnos libre y alegremente. Para utilizar el lenguaje de mis profesores, digamos que hemos mantenido con ellos una interacción no cooperativa con lo cual hemos obtenido un resultado subóptimo. Ahora podemos cambiar todo esto. Lo que tenemos que hacer es negociar con nuestros esclavos: los liberamos, desmantelamos todo el aparato coercitivo que implica la esclavitud y a cambio ellos se convierten en nuestros sirvientes voluntarios. Todos nos beneficiamos: nosotros obtenemos un mejor servicio, gastamos menos y ahorramos ese dinero para destinarlo a otros fines; ellos ya no reciben azotes, se liberan de sus cadenas y obtienen

mejores condiciones de vida, pues con el dinero que ahorramos al desechar la coerción y con el aumento de la productividad que nos deparará el servicio voluntario, podremos pagarles salarios en lugar de darles la pensión que les proporcionamos ahora... y hasta nos sobrarán recursos para agregar algunos *verdaderos* placeres a nuestras vidas. Ya me lo he imaginado todo; hay un profesor que ha desarrollado lo que llama el principio de concesión relativa minimáxima de las negociaciones racionales. Ese pacto con nuestros esclavos debe ajustarse perfectamente a tal principio. Ese profesor ha demostrado que es racional acatar voluntariamente tales pactos. De modo que no debe preocuparnos abandonar la coerción: nuestros esclavos son racionales, por lo tanto han de convertirse voluntariamente en nuestros sirvientes.

Los señores de más edad sacudieron la cabeza. La universidad no sólo costaba un montón de dinero, sino que además lo llenaba a uno de las ideas más condenadamente necias: semejante esquema hacía aguas por todos lados aun a los ojos de un miope. Quitemos las cadenas y jamás tendremos sirvientes voluntarios. Pero los jóvenes, impresionados por la perspectiva de la cooperación racional, no iban a amilanarse ante los argumentos de aquellos vejestorios que creían que el mejor modo de hacer la cosas era el adoptado por sus abuelos. Con la siguiente elección (porque los amos eran bastante democráticos entre sí) vio la luz la reforma administrativa y pocos años después la institución de la esclavitud quedó desmantelada, se dictó una proclama de emancipación y como broche de oro se inscribió en la constitución una solemne Negociación de Beneficio Mutuo.

¿Y todo ocurrió como lo había vaticinado el joven de nuestra anécdota? En absoluto. Los señores de más edad demostraron tener la razón. Como lo explicó uno de los ex esclavos, después de jurar como primer ministro de una administración destinada a derogar la Negociación de Beneficio Mutuo: "lo que esos jóvenes nunca comprendieron era que la negociación tenía una base coercitiva. Tal acuerdo sólo parecía racional a causa del poder que ellos ejercían sobre nosotros. Cuando ese poder perdió fuerza, resultó evidente que los frutos de la cooperación no se dividirían según ese fantasioso principio de la concesión relativa minimáxima. Y por lo tanto no había ninguna razón para

esperar que alguien lo acatara voluntariamente: nosotros no nos convertiríamos en sirvientes voluntarios. De todos modos, ellos se ahorraron una revolución, por lo tanto a fin de cuentas, ésta quizás haya sido la jugada más razonable que podían hacer, aun cuando la hayan hecho a causa de sus falsas expectativas".

Este relato nos sirve de fondo para desarrollar nuestro análisis de la posición negociadora inicial, o en realidad, de manera más general, de la posición inicial que tiene el actor tanto en la interacción de mercado como en la interacción cooperativa. En los capítulos 4 y 5 vimos que tanto la competencia del mercado como la cooperación acordada comienzan especificando los recursos disponibles y las expectativas de utilidad previas de las partes que interactúan. También establecimos la racionalidad y la justicia de procedimiento de tales formas de interacción. Pero señalamos explícitamente que los procedimientos justos sólo proporcionan un resultado imparcial si se parte de una posición inicial imparcial. Y es igualmente cierto que los procedimientos racionales sólo proporcionan un resultado racionalmente aceptable si se parte de una posición inicial racionalmente aceptable. En las observaciones que hace el primer ministro de nuestro relato admonitorio, está implícita la pretensión de que es racional acatar una negociación y que por lo tanto también es racional obrar de manera cooperativa sólo si la posición inicial no es coercitiva.

En la sección 2 desarrollaremos esta pretensión examinando dos enfoques que consienten la existencia de una coerción en la posición inicial y sosteniendo que tales enfoques no proporcionan una base para considerar racional el acatamiento. El primer enfoque identifica el resultado no cooperativo de la interacción natural con la posición negociadora inicial; el segundo identifica el punto de amenaza —el resultado que se obtiene si cada persona adopta una estrategia amenazadora de máxima efectividad— con la posición inicial. Negaremos que la conducta amenazadora sea relevante en el interacción racional y sostendremos que, si el resultado no cooperativo incluye la coerción, éste debe restringirse anulando los efectos de esa coerción, a fin de que pueda servir como una posición negociadora inicial para adoptar una estrategia conjunta que exija racionalmente el acatamiento individual.

Pero, el hecho de anular los efectos de la coerción de un resultado no cooperativo no nos brinda una caracterización positiva adecuada de la posición inicial. La idea principal que presentaremos y defenderemos en este capítulo es que si evaluamos el carácter racionalmente aceptable del resultado de la interacción desde el punto de vista de cada maximizador de la utilidad individual, o el carácter moralmente aceptable del resultado (es decir, que beneficia imparcialmente a cada individuo), comprobamos que es necesario anular de la posición inicial todos los efectos que pueda tener el hecho de sacar ventaja. Por lo tanto sostendremos que para cada individuo es racional y justo aceptar cierta restricción en la interacción natural y en la determinación de sus recursos disponibles iniciales, como una condición para ser voluntariamente aceptado por sus semejantes como una parte de los acuerdos cooperativos y de mercado, es decir de la interacción social. Esa restricción forma parte de una moral basada en el acuerdo, no por ser el objeto de un acuerdo entre individuos racionales, sino por ser una condición previa para que pueda realizarse tal acuerdo.

Al desarrollar esta argumentación, empezaremos caracterizando esa restricción y terminaremos demostrando que tiene fundamentos morales y racionales. En la sección 3 comenzamos pues estableciendo una forma de la estipulación lockeana que exige que "quede suficiente de algo igual y de bueno... para los demás"[1]. Mostramos cómo la estipulación obra como una restricción, pero en la sección 3 no mostramos que en realidad la estipulación determine los recursos disponibles iniciales, ni que sea imparcial ni que sea racional aceptarla. En la sección 4 encaramos las dos primeras de estas tareas y mostramos cómo aceptar la estipulación introduce en un estado de naturaleza previamente no moral una estructura de derechos; además sostenemos que esa estructura satisface la norma moral básica de imparcialidad. Y luego, en la sección 5 mostramos que para los maximizadores de la utilidad es racional aceptar la estipulación como una restricción de su interacción natural y de sus dotes individuales, en la medida en que intenten mantener una interacción social beneficiosa con sus semejantes.

Este último punto es esencial. Podemos decir que la estipulación moraliza y racionaliza el estado de naturaleza, pero sólo

en la medida en que concibamos el estado de naturaleza como una forma de llegar a la sociedad. Aunque para los seres humanos sea irracional permanecer en el estado natural sin aceptar restricciones a sus interacciones, ningún individuo puede beneficiarse restringiendo unilateralmente su conducta. Aceptar la estipulación equivale a acatar la primera ley de la naturaleza de Hobbes, "que todo hombre se esfuerce por alcanzar la Paz, siempre que tenga la esperanza de obtenerla".[2] Pero sin esa esperanza de pasar del estado natural a la sociedad, todo hombre puede pues "buscar y emplear todas las ayudas y las ventajas de la Guerra". Sin la perspectiva de un acuerdo y de una sociedad, no habría moral y la estipulación no sería racional. Afortunadamente, la perspectiva de una sociedad existe; lo que nos interesa es pues comprender la racionalidad de la moral que la sustenta.

2.1 En un convincente esbozo de una teoría contractual de la interacción racional[3] James M.Buchanan defendió la identificación de la posición negociadora inicial con el resultado no cooperativo. Puesto que coincidimos con muchos aspectos del enfoque de Buchanan debemos considerar cuidadosamente su argumentación.

Por cuestiones de simplicidad, Buchanan supone un mundo de dos personas con una mercancía escasa. Los habitantes de ese mundo no producen el bien sino que éste "les cae del cielo" en cantidades fijas. Esas personas no pueden beneficiarse comerciando entre sí ni invirtiendo en la producción. Sin embargo, podemos suponer que existe cierta interacción entre ellos, dictada por el deseo que cada uno abriga de consumir la mayor cantidad posible de la mercancía disponible. Cada uno de ellos puede tener razones, o puede creer tenerlas, para aplicar sus esfuerzos a obtener cierta porción del bien que originalmente cayó en manos del otro. En otras palabras, uno de los dos o ambos pueden invertir en la depredación. Y esa inversión puede darle a uno o a ambos razones para hacer una contrainversión en defensa. Finalmente, el resultado de esa interacción de depredación y defensa ha de ser lo que Buchanan llama la distribución natural, una condición de estabilidad en la cual la mezcla depredadora-defensiva de cada individuo es su respuesta maximizadora de la utilidad a la mezcla del otro. En nuestra

Pero, el hecho de
resultado no coopera
positiva adecuada de
presentaremos y defen
mos el carácter racion
interacción desde el p
utilidad individual, o
resultado (es decir, que
duo), comprobamos que
todos los efectos que pued
lo tanto sostendremos que
justo aceptar cierta restricción
determinación de sus recurso
condición para ser voluntariam
tes como una parte de los acuer
es decir de la interacción social.
una moral basada en el acuerdo
acuerdo entre individuos racional
previa para que pueda realizarse

Al desarrollar esta argumenta
rizando esa restricción y terminare
fundamentos morales y racionales.
pues estableciendo una forma de la
exige que "quede suficiente de algo
demás"[1]. Mostramos cómo la estipul
tricción, pero en la sección 3 no mostr
estipulación determine los recursos dis
sea imparcial ni que sea racional acep
encaramos las dos primeras de estas tar
aceptar la estipulación introduce en un
previamente no moral una estructura
sostenemos que esa estructura satisface la
de imparcialidad. Y luego, en la sección 5
los maximizadores de la utilidad es racional
lación como una restricción de su interacción
dotes individuales, en la medida en que inter
interacción social beneficiosa con sus semejantes

Este último punto es esencial. Podemos decir
lación moraliza y racionaliza el estado de naturaleza

terminología, esta distribución natural es el resultado no cooperativo.[4]

Pero ahora contamos con una base para el acuerdo. La distribución natural "sirve para establecer una identificación, una definición, de las personas individuales *partiendo de las cuales* es posible realizar acuerdos contractuales. Si no existiera ese punto de partida, sencillamente no habría manera de iniciar contratos significativos, ni en el plano real ni en el plano conceptual".[5] Podemos decir que la distribución natural le proporciona a cada persona una dote negociadora explícita; determina lo que cada individuo puede aportar a la mesa de negociación y por consiguiente constituye una posición negociadora inicial. Evidentemente, la mera existencia de una distribución natural no basta para que tengan lugar los contratos. Esa distribución debe ser subóptima, es decir que debe existir la posibilidad de un mejoramiento mutuo. Pero difícilmente pueda ponerse en duda ese carácter subóptimo ya que el esfuerzo dedicado a la depredación y la defensa es sumamente costoso. Ambas partes prefieren beneficiarse con un acuerdo que les permita apartarse de esas actividades improductivas. (Si alguien señalara que en los simplificados supuestos de Buchanan no hay lugar para la producción, podríamos responder que en cambio hay lugar para el ocio.)

Supongamos pues que se ha alcanzado el acuerdo partiendo de la distribución natural y supongamos también que, siguiendo nuestra argumentación del capítulo 5, ese acuerdo se basó en el principio de concesión relativa minimáxima. (Entre las preocupaciones de Buchanan no se halla la búsqueda de un principio de acuerdo.) De modo que cada persona ha de recibir su parte de la distribución natural o su equivalente en utilidad más una proporción del beneficio potencial que le brindaría la cooperación, en general igual al de la otra persona. ¿Es esto racionalmente aceptable para cada parte?

Seguramente la respuesta es no. En el ejemplo de Buchanan el carácter subóptimo de la distribución natural resulta de la inversión en depredación y la contrainversión en defensa. Mediante estas actividades cada individuo le impone al otro ciertos costos. Aunque la depredación pura, en la forma de la captura sin obstáculos, no es necesariamente costosa, el costo

que implica para el depredador la respuesta defensiva que provoca su depredación y el costo de esa defensa son improductivos. El acuerdo termina con esa actividad improductiva. Y alcanza un resultado óptimo en el cual están ausentes los esfuerzos depredadores/defensivos. Pero el efecto de esos esfuerzos continúa estando presente puesto que cada parte aporta a la mesa de la negociación los frutos de su actividad depredadora-defensiva y los obtiene (o sus equivalentes en utilidad) de esa mesa como parte del resultado general. Esos frutos no forman parte del superávit cooperativo que (en el ejemplo de Buchanan) está enteramente constituido por el consentimiento de cada parte de dejar de imponerle costos a la otra. Sólo esto cae bajo el influjo del principio de concesión relativa minimáxima.

Pero está claro que un individuo sería irracional si se dispusiera a acatar voluntariamente un acuerdo alcanzado así. Alguien que estuviera dispuesto a acatar los acuerdos que no modifican la distribución de los frutos de la depredación sencillamente estaría invitando a los demás a emprender actividades depredadoras y coercitivas como un preludio de la negociación. Estaría permitiéndoles a los depredadores que tuvieron éxito cosechar donde dejaron de sembrar, continuar beneficiándose de los efectos de la depredación natural después de haber participado de los acuerdos que los liberaron de la necesidad de invertir en nuevos esfuerzos depredadores. El acatamiento de la cooperación no equivale al acatamiento de la victimización. No negamos que, en la medida en que el costo de resistir la depredación real exceda cualquier beneficio que esa resistencia pueda proporcionar, racionalmente la víctima debe aceptar la depredación. Pero si la actividad depredadora queda proscrita, esa persona ya no tiene razones para obrar de un modo que mantenga sus efectos. Por consiguiente, el acuerdo alcanzado mediante la concesión relativa minimáxima de la distribución natural no provoca el acatamiento racional y voluntario de ambas (todas) las partes, si la distribución natural es en parte el resultado de la coerción.

El relato con que iniciamos este capítulo intenta ilustrar ese argumento. Los amos emplean la coerción a fin de mantener la obediencia de sus esclavos. La coerción es una práctica costosa para ambos. Tanto los amos como los esclavos se beneficiarían

si se desterrara la coerción y los esclavos continuaran sirviendo voluntariamente. Pero los ex esclavos no acatarían un acuerdo que tuviese ese efecto. Los esclavos proporcionan sus servicios porque los costos de la resistencia excede los beneficios que podría aportarles, teniendo en cuenta el poder de los amos. Pero lo único que los induce racionalmente a continuar dando servicio es el mantenimiento de ese poder. Sin la coerción, los ex esclavos aceptarían cierta forma de cooperación y se adherirían a ella, pero no a una cooperación basada en el resultado de la interacción coercitiva. Si fuera de otro modo, ¿por qué no pudo llegarse a un acuerdo del tipo propuesto por el joven de nuestro cuento entre los blancos y los negros de Sudáfrica? ¿O entre el gobierno polaco y los miembros de Solidaridad?

Buchanan es consciente de ese problema. Y se pregunta: ¿por qué habrían de acatar las personas voluntariamente las reglas y aceptar las instituciones de orden existentes?" cuando el acatamiento voluntario excluye aquellas basadas en la coerción y el castigo.[6] E insiste: "Sólo puede responderse a esta pregunta mediante una evaluación de la estructura existente, *como si* ésta fuera el resultado de un contrato admitido, o un contrato continuamente negociado. Los individuos deben preguntarse cómo son sus posiciones comparadas con aquellas que podrían haber esperado asegurarse en una situación contractual renegociada". Y lo que determina esas posiciones, al menos en parte, es considerar "los cambios imaginados de la distribución natural en el equilibrio anarquista que siempre subyace 'debajo' de las realidades sociales observadas".[7]

Buchanan supone que un individuo estaría dispuesto a acatar aquellas reglas (o, podríamos decir, aquella estrategia conjunta) que exigen que se establezca un acuerdo para que la distribución natural subyacente se haga realidad. Al decidir si es o no conveniente acatar, el individuo se pregunta: "Si el acuerdo caducara, ¿qué podría esperar yo?" Buchanan depende de la amenaza que está implícita en la distribución natural para conseguir el acatamiento. Pero, un regreso a la distribución natural no beneficia a nadie. La amenaza es irreal. Lo que motiva el acatamiento es la ausencia de la coerción antes que el miedo a que se renueve.

Consideremos nuevamente el sencillo mundo de dos perso-

nas que propone Buchanan y que representamos gráficamente en la figura 9. Un acuerdo que parta del punto I_n, la distribución natural o el resultado no cooperativo, conduciría a un resultado óptimo, el punto B_n. Pero B_n exige que el individuo V haga una trasferencia improductiva al individuo U que corresponda a la ganancia neta que U obtendría como depredador en la interacción natural. V tendría que darle a U, de manera completamente voluntaria, parte del bien que, en el ejemplo de Buchanan le cae del cielo. Hemos mostrado ya que en ausencia de la coerción V no tiene razón para hacer semejante cosa. Si se niega y conserva el bien para sí, el efecto es pues trasladar el resultado a los largo del límite o del borde superior derecho del espacio-resultado hacia B_c, el punto óptimo que representa el resultado del acuerdo logrado desde una posición inicial no coercitiva I_c. Si U objeta este resultado y trata de obligar a V a hacerle una transferencia, debe pues reintroducir la coerción y de ese modo el resultado se traslada a lo largo de la línea que une B_c con la distribución natural I_n, un traslado que es costoso para las dos partes y además irracional para U. De ahí que debamos suponer que un acuerdo estable termina en B_c, aun cuando las partes comiencen realmente en I_n. El resultado es el mismo al que finalmente se llegaría si I_c fuera la posición negociadora inicial.

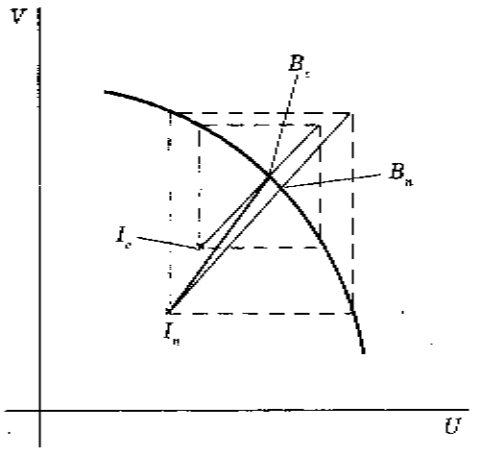


Figura 9

En el sencillo mundo de Buchanan, cada individuo ha de disfrutar simplemente del bien que le ha tocado en suerte y dedicar al ocio el esfuerzo que antes invertía en la depredación y la defensa. Pero en un mundo más complejo podemos suponer que ese esfuerzo puede dirigirse a un empleo productivo, de modo tal que las partes participen de acuerdos de mercado o de la producción cooperativa y obtengan bienes que antes les eran inaccesibles. Las ganancias que pueden esperarse del acuerdo aumentan pues considerablemente. Y así si *V* se niega a hacerle una transferencia improductiva a *U*, para tratar de forzarlo, éste debe no sólo reintroducir la coerción, sino también renunciar a los bienes logrados gracias al empleo productivo de los esfuerzos que permitió el acuerdo al dejar que cada individuo no tuviera que ocuparse ni de la depredación ni de la defensa. Se hace aun más evidente entonces el costo que tiene para *U* reintroducir la coerción y por lo tanto la irracionalidad de tratar de obligar a *V* a hacer una trasferencia improductiva.

A fin de vincular este análisis elemental con nuestra sociedad de amos y esclavos, observemos que el punto I_n corresponde a la situación en que se hallaban los amos y los esclavos al comienzo de nuestro relato. B_n representa el resultado del acuerdo según lo imaginaba ingenuamente el joven universitario, en el cual sin la coerción los ex esclavos continúan sirviendo a sus ex amos. La provisión de tales servicios es la transferencia improductiva que exigía el acuerdo de los ex esclavos, transferencia que, por supuesto, éstos rechazaron. Ese rechazo trasladó a la sociedad de B_n a B_c, el resultado al que se llegó después de la promulgación de la llamada Negociación de Beneficio Mutuo. Y ése es el resultado que se habría obtenido si el acuerdo se hubiese alcanzado desde una posición negociadora inicial no cooperativa pero también no coercitiva, representada por el punto I_c.

Hemos sostenido que un individuo debería considerar irracional disponerse a aceptar voluntariamente una estrategia conjunta acordada desde una posición inicial coercitiva. A la luz de nuestro análisis podemos reformular nuestra afirmación y decir que un individuo debe considerar irracional disponerse voluntariamente a hacer transferencias improductivas a los demás. Una transferencia improductiva no origina nuevos bienes ni implica intercambio alguno de los bienes existentes;

sencillamente redistribuye algunos bienes existentes de una persona a otra. Por consiguiente, implica un costo de la utilidad, a cambio del cual no se recibe ningún beneficio y una ganacia de utilidad a cambio de la cual no se brinda ningún servicio.

Sólo es racional hacer una transferencia improductiva si el hecho de hacerla es directamente maximizador de la utilidad. Puesto que la transferencia es en sí misma costosa, sólo puede ser maximizadora de la utilidad en la medida en que se la imponga coercitivamente. Por lo tanto, no puede formar parte de ninguna interacción cooperativa puesto que una interacción semejante incluye el beneficio mutuo. El hecho de que se hagan transferencias improductivas en otros acuerdos cooperativos prueba que existe un residuo de la depredación natural.

En nuestros ejemplos es fácil detectar la depredación residual. Pero en el mundo real es algo que se disfraza prontamente. Las transferencias improductivas se esconden detrás de adornos morales e ideológicos espurios. Si bien ya no nos inclinamos ante los halagos de la nobleza, todos estamos demasiado prontos a sucumbir a los pies de la inequidad. Pero las personas idealmente racionales no se dejan llevar por esto y reconocen las transferencias improductivas. Pueden ejercer y sufrir la coerción, pero no la confunden con la cooperación.

Puesto que el resultado de la estrategia cooperativa conjunta no está en equilibrio, carece de estabilidad natural. El acatamiento proporciona una estabilidad artificial pero más débil en los casos en que no es posible lograr la estabilidad natural más firme. Si tuviéramos que considerar la distribución natural como la posición negociadora inicial, en algunos casos comprobaríamos pues que las transferencias improductivas son necesarias para alcanzar el resultado cooperativo. Aun cuando dentro de los acuerdos cooperativos puedan imponerse coercitivamente ciertas transferencias particulares, los acuerdos mismos dependen en definitiva del acatamiento voluntario. Y este acatamiento podría no provenir de aquellos a quienes se les pide que hagan la transferencia. En ese caso el resultado cooperativo carecería de estabilidad tanto natural como artificial. El Necio de Hobbes podría mostrar que, aun cuando pueda parecer racional participar de la interacción cooperativa, puede no ser racional acatar sus exigencias.

Llegamos a la conclusión de que la posición negociadora

inicial, como punto de partida de una cooperación racional, puede no identificarse con la distribución natural o con el resultado no cooperativo. La distribución natural respresenta los efectos del poder. Ahora bien, no hemos mostrado, ni mostraremos, ninguna base, racional o moral, que permita criticar los efectos del poder considerados en sí mismos. Pero, si consideramos la distribución natural en relación con el mercado o con la interacción cooperativa, la evaluamos no en sí misma pero sí en cuanto a su capacidad para determinar lo que cada persona aporta al mercado o a la negociación implícita en los arreglos cooperativos. Estimamos la distribución natural como aquello que determina las dotes de cada persona. Y aquí encontramos una base para criticarla y en realidad para rechazarla, pues en el fondo es coercitiva. Si consideramos que las dotes de cada persona es la que determina sus derechos (una visión que desarrollaremos en 4.1 de este capítulo), podemos pues decir que aceptar la distribución natural como la posición negociadora inicial equivaldría a aceptar que el poder determina los derechos.[8] Es indudable que a veces se invoca el derecho para mantener lo que estableció el poder. Pero el derecho así invocado es un impostor, incapaz de pasar el examen de la racionalidad maximizadora de la utilidad.

2.2 En la teoría de la negociación racional desarrollada por John Nash y generalizada por John Harsanyi, se identifica la posición negociadora inicial con el *punto de amenaza* que representa el resultado que se alcanzaría si cada persona obrara según su estrategia amenazadora máximamente efectiva.[9] ¿Cuáles son esas estrategias?

En 5.3.4 vimos que desde el punto de vista de Zeuthen-Nash-Harsanyi, si la posición inicial de dos personas, Ann y Adam, está representada por el punto (u^*, v^*), el resultado debe estar pues representado por el punto (u', v') tal que para cualquier punto (u, v) del espacio-resultado, el producto $(u' - u^*)(v' - v^*)$ es por lo menos tan grande como el producto $(u - u^*)(v - v^*)$. De este modo podemos correlacionar cada punto del espacio-resultado, considerado como una posición negociadora inicial, con un punto situado en el límite superior derecho del espacio-resultado, como el resultado negociado correspondiente a esa posición inicial.

Inversamente, podemos correlacionar cada punto situado en el límite superior derecho del espacio-resultado con ese conjunto de puntos del espacio-resultado, cada uno de los cuales (considerado como posición negociadora inicial) lo determina como el resultado de la negociación. Está claro que cada punto del espacio-resultado corresponde a uno y sólo uno de tales conjuntos; dos puntos pertenecientes al mismo conjunto determinan el mismo resultado; dos puntos pertenecientes a diferentes conjuntos determinan diferentes resultados. Llamaremos a estos conjuntos *conjuntos de resultado equivalente.*

Está claro que al elegir un punto que sirva de posición negociadora inicial, tanto Ann como Adam son indiferentes entre todos los puntos pertenecientes al mismo conjunto de resultado equivalente, pero tienen preferencias estrictamente opuestas entre dos puntos cualesquiera pertenecientes a conjuntos diferentes, que corresponden a sus preferencias estrictamente opuestas entre dos resultados óptimos cualesquiera. Al elegir un punto que les sirva de posición negociadora inicial, Ann y Adam se interesan por, y sólo por, el conjunto de resultado equivalente al cual corresponde aquel punto.

Supongamos que aunque las estrategias de que disponen Ann y Adam permanezcan fijas, los rendimientos de los resultados cambiaran, de manera tal que lo que llamaremos la utilidad *e* de cada resultado sea una medida de preferencia para el conjunto de resultado equivalente al que pertenece. (La preferencia por cada conjunto fue por supuesto determinada por la preferencia por su resultado negociado asociado.) Entonces resulta que Ann y Adam afrontan una situación estrictamente competitiva, en la cual cada uno tiene una estrategia maximizadora de la utilidad *e* contra la estrategia del otro, y que por lo tanto lo lleva a obtener un resultado en equilibrio *e*, que también es óptimo *e*. Estas son sus estrategias amenazadoras máximamente efectivas; si Ann y Adam obraran con la intención de determinar una posición negociadora inicial, racionalmente elegirían estas estrategias.

Este análisis muy esquemático no puede aplicarse directamente a solucionar nuestro problema de la negociación, pero afortundamente no necesitamos afrontar el problema de determinar las estrategias amenazadoras de máxima efectividad.

Porque, consideremos por un momento qué significan esas estrategias. En el análisis de Nash-Harsanyi cumplen una función puramente hipotética, puesto que Ann y Adam en realidad no las eligen, sino que meramente apelan a ellas para determinar los costos que cada uno puede imponerle al otro en una competencia estricta para lograr una ventaja en la negociación. Si Ann y Adam vieran que son incapaces de cooperar entre sí tampoco elegirían las estrategias amenazadoras máximamente efectivas; el punto de amenaza no tiene una particular relación con el resultado no cooperativo. Pero si Ann y Adam no pudieran elegir esas estrategias, no podrían amenazarse de manera creíble entre sí. Las estrategias amenazadoras máximamente efectivas demuestran ser vanas.

Generalmente los teóricos de la negociación suponen que los individuos están en posición de hacer cumplir sus amenazas.[10] Pero, en la mayor parte de las situaciones, ésa es una suposición irreal. Y ya hemos visto que en la cooperación la conducta amenazadora ha de ser proscrita. Por lo tanto, aun dejando de lado la irracionalidad que implica acatar voluntariamente un acuerdo que toma el punto de amenaza como la posición negociadora inicial, llegamos a la conclusión de que los cooperadores potenciales no se disponen a ejercer la coerción mediante la amenaza. De modo que no vale la pena agregar nada más referente al punto de amenaza.

3.1 La posición negociadora inicial debe ser no coercitiva. Pero, ¿debemos avanzar más en restringir la interacción natural teniendo en cuenta que ésta determina la base del mercado o de la interacción cooperativa? Sostendremos que los términos de una cooperación plenamente racional incluyen la exigencia de que los recursos de cada individuo, que le proporcionan una utilidad básica no incluida en el superávit cooperativo, debe considerarse como inicialmente adquirida por ese individuo sin que éste haya sacado ventaja de ningún otro o, más precisamente, de ningún otro cooperador. De lo contrario, aquellos que consideran que estuvieron en desventaja en la adquisición inicial han de considerar que la sociedad es injusta al exigirles pagos sin ofrecerles a cambio una compensación, y no han de tener razones suficientes para aceptar los arreglos del mercado o para adherirse voluntariamente a las estrategias cooperativas conjuntas.

En esta sección y en gran parte de la siguiente trataremos de comprender la exigencia. Consideraremos qué significa sacar ventaja y mostraremos cómo el hecho de no sacar ventaja restringe la adquisición de recursos para la interacción del mercado o la interacción cooperativa. En nuestro análisis trataremos de darle un carácter plausible a la idea de que no sacar ventaja es una restricción razonable y justa que la interacción natural debe satisfacer en la medida en que su resultado proporcione una posición inicial para la negociación. Entonces comprenderemos que no sacar ventaja es una manera de excluir a los jugadores líberos y a los parásitos, de modo tal que al ser una restricción amplía al estado natural la imparcialidad que logran el mercado y la cooperación. Pero que algo sea plausible no lo convierte en una prueba; después de examinar los efectos que tiene en la interacción natural no sacar ventaja de los demás, trataremos de demostrar que no sacar ventaja es una exigencia moral y racional. Y una vez más debemos destacar que nada de nuestro análisis implica que sería racional refrenar los deseos de sacar ventaja de los demás o acatar cualquier otra restricción a la maximización directa, desde el estricto punto de vista de la interacción que se desarrolla en un estado de naturaleza. Permanentemente concebimos el hecho de no sacar ventaja como algo que restringe la interacción natural con el único fin de facilitar la aparición de la sociedad.

El primer gran intento de racionalizar y moralizar el estado de naturaleza hobbesiano —esa condición de depredación ilimitada— es la teoría de la propiedad de John Locke.[11] El principio de adquisición que constituye la médula de esa teoría sugiere una formulación preliminar de la exigencia de no sacar ventaja. Pero ampliaremos la posición de Locke de modo tal de poder expresar su restricción sobre la adquisición como una estipulación que autoriza y limita simultáneamente los derechos exclusivos de los individuos sobre objetos y poderes. El efecto de tal restricción es proporcionarle a cada persona una esfera de control exclusivo al prohibirles a los demás interferir con algunas de sus actividades. Esa esfera exclusiva constituye una esfera moral que define al individuo en sus relaciones de mercado y en sus relaciones cooperativas.

Según Locke una persona adquiere un derecho exclusivo al

producto de su trabajo, con la condición de que lo utilice o al menos no lo malgaste y siempre que deje para los demás "suficiente de algo igual de bueno", "más de lo que podrían servirse los aún desprovistos".[12] Por lo tanto, la posición negociadora inicial que determina este principio implica una distribución de factores de producción y de otros bienes basada en el trabajo natural de cada persona, en la medida en que ese trabajo se relacione con el uso, y en la medida en que la adquisición de cada persona deje un excedente. El trabajo, el uso y el excedente son los tres conceptos clave que permiten comprender el punto de vista de Locke y cada uno de ellos parece plantear profundos problemas.

La teoría de Locke de la adquisición moraliza el estado de naturaleza hobbesiano. Pero ¿cómo es posible? ¿Cómo puede mi trabajo, al cual según Hobbes tengo libertad, constituir la base de mis derechos? ¿Cómo puede mostrar Locke que tengo un derecho a mi cuerpo y sus poderes, salvo suponiéndolo y eludiendo así la cuestión de establecer una posición negociadora racional e imparcial? A fin de resolver estas dificultades debemos enfocar la idea de la adquisición original o inicial atendiendo a la restricción del excedente, lo cual, según comprobaremos, constituye su aspecto fundamental. La adquisición del propio cuerpo y sus poderes, la función que cumple el trabajo y la exigencia de que cada uno utilice lo que adquiere, cobran su verdadero sentido después de una conveniente interpretación de la estipulación lockeana expresada en las palabras "suficiente y tan bueno", una estipulación que restringe las interacciones naturales a fin de que la sociedad sea posible.

Tomar literalmente estas palabras podría parecer una forma de aniquilar nuestra empresa desde el comienzo. Si uno debe dejar lo suficiente de igual bondad para los demás, ¿puede acaso pretender algo en absoluto como recurso disponible? Hasta la adquisición del propio cuerpo parece invalidada porque si una persona es más capaz y más fuerte que yo, la pretensión de su cuerpo y sus poderes no dejaría lo suficiente ni tan bueno para mí que soy más débil y menos capaz. A mí me convendría compartir a medias las facultades conjuntas del primero que quedarme con todas las mías y ninguna del más fuerte. Por lo demás, exigir que quede lo suficiente y tan bueno seguramente

impediría la adquisición no sólo de lo que realmente escasea, sino también de aquello de lo que hay potencialmente una oferta insuficiente. No puede haber suficiente de lo que escasea, de modo tal que si una persona toma lo suficiente para sí, lo que queda para los demás no es tan bueno. Esta interpretación literal de la estipulación lockeana sencillamente no nos da una definición de las personas relacionada con los fines de la negociación o del mercado. El mundo y todo lo que hay en él sería y continuaría siendo pasto común.

Por lo tanto, debemos recurrir a la interpretación de la estipulación lockeana que ofreció Robert Nozick. "La estipulación que establece Locke *de que se deje lo suficiente y tan bueno en común para los demás* equivale a asegurar que no se empeore la situación de los demás."[13] Si la interacción natural está destinada a determinar la posición negociadora inicial, debe excluir actividades que empeoren la situación de cualquier persona, ya sea mediante la depredación o por otros medios. Las dotes de cada persona incluyen todo aquello que esa persona adquiera sin empeorar la situación de sus semejantes; esas dotes le proporcionan pues una base para aplicar la concesión relativa minimáxima y determinar así el resultado esperado de la cooperación. Cada individuo es libre de utilizar los recursos de sus dotes para aumentar lo que aporta a la mesa de la negociación y también, por supuesto, para participar en la competencia del mercado.

Pero prohibir sencillamente que empeore la situación de los demás es demasiado fuerte. Porque hay situaciones en las cuales una persona sólo puede impedir eso empeorando su propia posición. Siguiendo a Locke, quien permite que se dé justificable prioridad a la propia preservación antes que a la de los demás en las deliberaciones que uno hace, modificamos la interpretación de la estipulación que da Nozick y decimos que prohíbe empeorar la situación de los demás excepto en los casos en que sea necesario para impedir que empeore nuestra propia situación. Está claro que cualquier otra estipulación más firme fracasaría como restricción posible para determinar la posición negociadora inicial entre maximizadores de la utilidad.

¿Cómo debemos entender lo que significa empeorar —e inversamente, mejorar— la situación de alguien cuando ese

alguien es uno mismo y cuando es otra persona? No hay problemas para establecer qué es "mejor" y qué es "peor"; una situación es mejor para una persona que para otra, si y sólo si le proporciona una mayor utilidad esperada. Ahora bien, la situación de un determinado individuo mejora o empeora sólo en relación con algún punto de base. Ese punto de base no puede ser el mismo de la posición negociadora inicial, puesto que esa posición debe quedar determinada por la aplicación de la estipulación que prohíbe empeorar la situación de otro salvo si es necesario para no empeorar la propia. Debemos preguntarnos qué aspecto de la interacción natural puede ser tomado como punto de base para determinar qué efectos producen las acciones de una persona sobre la situación de otra o sobre su propia situación. La distinción crucial que debemos establecer es entre empeorar la situación de alguien y no mejorarla, puesto que la estipulación prohíbe sólo lo primero y no lo último.

A fin de ilustrar nuestra indagación en este punto, consideremos un ejemplo caro a los filósofos. Una persona se está ahogando en el río y otra que pasa por la orilla la deja ahogar. Este es un resultado; consideremos dos maneras en las que se pudo llegar a él. Primero la persona cayó al agua. La otra pasa, oye los gritos de auxilio pero los ignora y continúa su camino. La segunda posibilidad: la persona está parada en la orilla. El otro llega, lo empuja al río e ignorando los pedidos de auxilio continúa su camino. En el primer caso, aunque el que pasa no mejora la desgraciada situación del que se ahoga, tampoco la empeora. En el segundo caso, aunque el resultado sea el mismo, el segundo personaje empeoró claramente la situación del primero. ¿Por qué esa diferencia?

Supongamos que la segunda persona no hubiese pasado por allí. En ese caso lo más probable es que la primera se ahogara; y en el segundo caso hubiera permanecido a salvo en la orilla. En el primer caso el resultado para el primer personaje no es peor del que hubiera obtenido en ausencia del segundo personaje, de ahí que no se puede decir que éste haya empeorado la situación de aquél. Sin embargo, teniendo en cuenta que el transeúnte podría haberlo salvado o ayudado a salvarse, el resultado en el primer caso podría haber sido mejor que el que se obtendría en ausencia del transeúnte, si éste hubiera obrado

de manera diferente; de ahí que el segundo personaje no mejoró la situación del primero. (Si, por la razón que fuera, el transeúnte hubiese sido incapaz de hacer algo para ayudar al que se ahogaba, tampoco hubiese mejorado la situación pero no podríamos decir que fracasó tratando de mejorarla.) En el segundo caso, el resultado es para el primer personaje peor de lo que hubiera sido en ausencia del segundo; de ahí que podamos decir que el transeúnte empeoró la situación de aquél.

En este ejemplo sencillo, el punto de base para establecer cómo el segundo personaje afecta al primero, en cuanto a mejorar o empeorar su situación, queda determinado por el resultado que el primero podría esperar en ausencia del segundo. El empeoramiento e igualmente el mejoramiento pueden juzgarse comparando lo que realmente hizo el segundo personaje con lo que podría haber ocurrido, *ceteris paribus*, en su ausencia. No hacer nada por mejorar e igualmente no hacer nada por empeorar pueden juzgarse comparando lo que podría haber hecho el segundo personaje y no hizo con lo que podría haber ocurrido en su ausencia. El transeúnte empujó al primer hombre que en su ausencia no hubiese caído; es decir que empeoró su situación. El transeúnte podría haber salvado al primer hombre; pero éste no se habría salvado en ausencia de aquél; entonces el transeúnte no hizo nada por mejorar la situación del primero.

Es complicado determinar el punto de base apropiado que este ejemplo establece. Si una persona se está ahogando y yo soy un nadador experto empleado para proteger la vida de los bañistas, al ignorar los gritos de auxilio de esa persona estoy ayudando a empeorar su situación. Porque aquí mi conducta debe juzgarse comparando con lo que habría ocurrido en mi ausencia en una playa normalmente protegida y no en una playa en la cual no esté yo ni ningún otro cuidador. Pero obsérvese que el hecho de que no haya un guardavidas, suponiendo que no haya ninguna exigencia ni compromiso previo, no puede empeorar o contribuir a empeorar la situación de los bañistas, aun cuando, una vez que se contrata a uno el hecho de que éste desatienda su deber puede empeorarla. Este tipo de complicaciones indica la importancia de considerar el marco institucional y la naturaleza de las prácticas dentro de las cuales

se dan las acciones, a fin de determinar el apropiado punto de base para juzgar qué significa empeorar o mejorar una situación.

Hasta aquí sólo hemos considerado el efecto que puede producir una persona sobre la situación de otra, pero ¿qué ocurre si una persona afecta su propia situación? En ese caso no podemos apelar a su ausencia a fin de determinar el punto de base. Sin embargo, el supuesto de interacción proporciona un contexto determinado para mejorar o empeorar la propia situación. Aunque podemos hablar de que alguien mejora su situación en la medida en que prefiere el resultado obtenido tanto a su posición previa como a otro resultado alternativo que podría haber alcanzado, esto no es relevante para el presente análisis. Obsérvese en cambio que así como yo mejoro la situación de una persona en la medida en que esa persona prefiera el resultado de la interacción conmigo a lo que podría haber esperado en mi ausencia, del mismo modo puedo mejorar mi propia situación en la medida en que yo prefiera el resultado de la interacción con otra persona al que podría haber esperado si esa persona no existiera o no estuviera disponible para la interacción. Y empeoro mi propia situación estando en interacción con otra persona si prefiero el resultado que podría haber esperado en su ausencia al que realmente voy a obtener de la interacción. Una vez más el ejemplo prescinde de cualquier marco institucional.

Interpretamos la estipulación lockeana de modo tal que entendemos que ésta prohíbe empeorar la situación de otra persona, salvo si se lo hace para impedir el propio empeoramiento mediante la interacción con esa persona. O, podemos decir, la estipulación prohíbe mejorar la propia posición mediante una interacción que empeore la situación de otro. Esta manera de decirlo nos parece que expresa más claramente la idea implícita en la estipulación de Locke sobre no sacar ventaja.

3.2 La estipulación lockeana se aplica a la interacción supuestamente regida por la racionalidad maximizadora de la utilidad individual y el desinterés mutuo. Se supone que cada persona elige una estrategia que maximiza su utilidad esperada, salvo que la estipulación prohíba expresamente hacerlo. Cada individuo es pues libre de mejorar su propia situación como prefiera, con la condición de que al hacerlo no empeore la

situación de otra persona. Exigir que, como una condición de mejorar la propia situación, la persona debe mejorar la de los demás, equivaldría a exigir que se permita la existencia de juegos libres. Pero nadie es libre de mejorar su propia situación mediante una interacción que empeore la situación de otro. Permitir que a fin de mejorar la propia situación una persona pueda empeorar la de los demás equivaldría a permitirle a esa persona que sea un parásito. De modo tal que aplicar una restricción más firme a la interacción natural que la estipulación lockeana implicaría permitir la existencia de los jugadores líberos y aplicar una restricción más débil sería autorizar el parasitismo. Navegando entre esas dos aguas, la estipulación restringe la interacción natural a fin de hacer posible una cooperación racional, justa y libre.

Puede suponerse que una persona que acepta la estipulación como una restricción a su conducta maximizadora razona acerca de la elección de una estrategia del modo siguiente:

> (1) Divide sus estrategias en tres grupos:
> *A*: aquellas que le proporcionan a cada persona una utilidad esperada no menor que la que podría esperar en ausencia de la interacción;
> *B*: aquellas que le proporcionan a ella una utilidad no menor que la que podría esperar en ausencia de la interacción, pero que le brindan a alguna otra persona una utilidad esperada menor que la que esa persona podría esperar en ausencia de la interacción;
> *C*: aquellas que le proporcionan a ella una utilidad esperada menor que la que podría esperar en ausencia de la interacción.
> (2) Si el Grupo *A* no es un conjunto vacío, la persona elige pues de allí una estrategia que maximice su utilidad esperada. Mejora, o por lo menos no empeora su situación, sin empeorar la de ningún otro individuo.
> (3) Si el Grupo *A* es un conjunto vacío y el Grupo *B* no lo es, la persona elige pues del Grupo *B* una estrategia que minimice las pérdidas que han de sufrir los demás en comparación con la utilidad que podrían esperar en ausencia de la interacción. La persona no empeora su propia situación y si es compatible con esto, minimiza el empeoramiento de la situación de los demás. La estipulación prohíbe el empeoramiento innecesario.
> (4) Si los Grupos *A* y *B* son conjuntos vacíos, la persona elige pues del Grupo *C* una estrategia que maximice su utilidad esperada. Minimiza el empeoramiento de su propia situación cuando la interacción no le es provechosa.

Consideremos un ejemplo de interacción restringida por la estipulación. Joanna y Jonathan son dos náufragos perdidos en una isla pequeña e inhabitada. Interesados en establecer las bases para la cooperación, cada uno de ellos decide acatar la estipulación. Por consiguiente, Jonathan no obliga a Joanna a someterse a sus deseos sexuales aun cuando eso podría mejorar su situación (en comparación con la ausencia de la mujer), porque podría empeorar la situación de ella (en comparación con la ausencia del hombre). Pero cuando Joanna desea también mantener relaciones sexuales, Jonathan maximiza su satisfacción sin preocuparse por el placer de la mujer. El mejora su situación sin empeorar la de Joanna. Luego, por supuesto, ambos pueden acordar aumentar el placer sexual del otro y comprobar que eso es mutuamente beneficioso. Pero ese mejoramiento mutuo es cooperativo.

Cada uno toma el alimento que desea, sin preocuparse por el otro. Puesto que en ausencia del otro cada uno podría tener acceso a todo el alimento que proporciona la isla, cada uno podría empeorar su situación si frenara su impulso de tomar todo el alimento que desea y puede obtener. Sin embargo, cuando Joanna comienza a cultivar una huerta y cosecha mayor cantidad de frutas y verduras, Jonathan refrena su impulso de tomar lo que desea porque si lo hiciera estaría mejorando su situación (pues en ausencia de la mujer no habría huerta) y empeorando la de Joanna (pues en ausencia del hombre ella podría disfrutar de la huerta sola). Y cuando Jonathan confecciona una red de pesca y captura una mayor cantidad de peces, Joanna, por la misma razón, refrena su impulso de tomar todo el pescado que desea.

Por supuesto, para entonces es lógico esperar que Joanna y Jonathan hayan alcanzado un acuerdo de cooperación —y algo más si tenemos en cuenta la pronunciada curva que ha adoptado el vientre de Joanna—. Pero sólo un acuerdo semejante —explícito o implícito— puede conducirlos a un estado de cosas óptimo. En la interacción natural, aun restringida por la estipulación, cada individuo evalúa cada situación de manera separada al decidir qué hacer, de modo tal que sólo puede tenerse en cuenta la reciprocidad inmediata. En la cooperación, los individuos pueden extender su horizonte de interés y aceptar ahora costos de acciones que han de beneficiarlos luego.

En la interacción natural sería irracional para Joanna correr cualquier riesgo destinado a salvar a Jonathan de los tiburones (suponiendo, por supuesto, el desinterés mutuo). Si Joanna empeorara su situación de ese modo mejorando la de Jonathan, estaría permitiéndole ser un parásito. Pero si han convenido cooperar probablemente se encuentren con la obligación moral y racional de hacer esfuerzos por salvar al otro aun corriendo riesgos personales. Tanto Joanna como Jonathan han de beneficiarse de una práctica que les imponga a ambos ciertos costos a fin de conferirle al otro mayores beneficios. Al convenir el mutuo salvamento, cada uno se asegura que el otro no ha de dejarlo librado a cualquier peligro que pueda presentarse.

La estipulación no incorpora pues en el estado de naturaleza las demandas de optimidad justa. Meramente restringe la interacción natural a fin de determinar una posición inicial partiendo de la cual pueda alcanzarse un resultado óptimo y justo. O eso es lo que pretendemos; pero recordemos que aun tenemos que mostrar que la estipulación satisface las normas de la moral y la razón.

4.1 Locke presenta la estipulación para restringir una actividad particular, la adquisición o apropiación de objetos externos. Y sostiene que el ejercicio de esa actividad de conformidad con la estipulación le proporciona al actor un derecho exclusivo a la tierra o a los bienes de los que se apropió. De ese modo el actor obtiene las dotes materiales en los recursos de producción necesaria para definirlo plenamente a los fines de la interacción del mercado y la interacción cooperativa. Pero Locke supone que cada persona comienza con un derecho exclusivo a su cuerpo y sus poderes, un derecho que se extiende, mediante el trabajo, a aquello que la persona adquiere.

Para nosotros la estipulación cumple una función fundamental más amplia. La consideramos una restricción general, mediante la cual podemos pasar del estado natural hobbesiano, en el cual no existen los derechos exclusivos, sino sólo las libertades, a la posición inicial necesaria para desarrollar la interacción social. Comenzamos por considerar lo que cada persona puede hacer con ese cuerpo y sus poderes a los cuales tiene acceso directo, y qué pueden o no hacer los demás con ese cuerpo y sus poderes. Mostramos que cada persona tiene ciertos

derechos a su cuerpo y sus poderes y que esos derechos constituyen su dote básica según la definimos en 4.3.2. Luego consideramos cómo esos derechos, que Locke supone y nosotros justificamos, conducen a adquirir nuevos derechos sobre la tierra y otros bienes. Esos nuevos derechos surgen no sencillamente de actos de apropiación conformes a la estipulación, sino también de lo que una persona puede hacer y de lo que en consecuencia los otros no pueden hacer, en suma, del modo en que un individuo puede utilizar ciertos bienes y las consecuentes restricciones que les impiden a los demás usar esos bienes e interferir en el uso que el primero quiera darles.

De este modo, mostramos que la estipulación introduce en la interacción una estructura rudimentaria de derechos. Convierte la condición natural depredadora descrita por Hobbes en la condición natural productiva supuesta por Locke. Pero su función primaria es posibilitar la aparición de las estructuras necesarias para dar forma a la interacción social tanto cooperativa como competitiva. Evidentemente podemos imaginar otros modos posibles de dar forma a la interacción social, otras maneras de establecer una estructura de derechos que determine las dotes iniciales de cada persona. Lo que decimos es que la estipulación lo hace de un modo racional e imparcial.

En el estado natural de Hobbes un individuo puede usar su cuerpo y sus fuerzas como mejor le plazca. Aunque la estipulación restringe ese uso, lo hace sin afectar el fundamento de la libertad manifestada. Porque al ejercer los propios poderes una persona no necesita entrar en interacción con otras y por lo tanto tampoco necesita mejorar su situación en relación con lo que podría esperar en ausencia de los demás, ni tampoco empeorar la situación de éstos en relación con lo que ellos podrían esperar en ausencia de aquella. Cualquier restricción derivada de la estipulación afecta el modo en que un individuo usa sus propios poderes pero no el mero hecho de utilizarlos.

En el estado de naturaleza de Hobbes un individuo puede utilizar también los cuerpos y las fuerzas de otros como mejor le plazca; dentro de los límites de lo que ese individuo es capaz de hacer no hay nada que no le esté permitido. Pero la estipulación afecta directamente esa libertad. Porque al utilizar los poderes o fuerzas de otro ese individuo mejora o espera mejorar su propia

situación mediante la interacción con ellos, y en la medida en que el uso que haga esa persona deba interferir con el ejercicio que los demás quieran hacer de sus poderes, empeora sus situaciones mediante esa interacción.

En ausencia de sus semejantes cada persona puede esperar utilizar sus propios poderes pero no los de los otros. Esta diferencia es crucial. Porque nos brinda el punto de base partiendo del cual puede aplicarse la estipulación a la interacción. El uso continuado de los propios poderes en presencia de los demás no mejora en sí mismo la situación de uno; el uso de los poderes de los demás sí mejora la situación del individuo. Frenar el impulso de utilizar los propios poderes empeora la situación de una persona; frenar el impulso de utilizar los poderes de los otros no logra mejorar la situación de ese individuo pero tampoco la empeora. El uso continuo de los propios poderes quizá no logre mejorar la situación de los demás pero en sí mismo tampoco la empeora; el uso de los poderes ajenos, es decir interferir en el uso que los otros puedan hacer de sus propios poderes, sí empeora la situación de los demás. Por lo tanto, al prohibir que cada individuo mejore su situación empeorando la de los demás, pero dejando que aparte de eso cada uno haga lo que mejor le plazca, la estipulación no sólo le confirma a cada uno el uso de sus poderes, sino que además, al impedirles a los otros que los usen, le proporciona a cada individuo el uso exclusivo de sus propios poderes. De modo que la estipulación convierte las libertades ilimitadas de la naturaleza de Hobbes en derechos y deberes exclusivos. Cada persona tiene un derecho exclusivo para ejercer sus propios poderes sin que los demás le opongan obstáculos y el deber de refrenar el impulso de utilizar los poderes ajenos en la medida en que ese uso interfiera en el ejercicio que los demás puedan hacer de los poderes a los que tienen acceso directo.

Hemos establecido así la justificación de las *dotes básicas* prometida en 4.3.2. Suponemos que cada persona se identifica con aquellas aptitudes físicas y mentales a las que tiene acceso directo y vemos que esa identificación le proporciona a cada persona un sentido normativo de sí misma, expresado por su derecho a utilizar tales aptitudes. Al apelar a la estipulación mostramos que esa identificación no es arbitraria, sino que

antes bien está plenamente justificada por la racionalidad y la imparcialidad de la estipulación (que aún debemos demostrar).

Hemos dado ya el primer paso de la conversión del estado de naturaleza; aplicar la estipulación le da a cada persona el derecho exclusivo a utilizar su cuerpo y sus poderes, sus aptitudes físicas y mentales. El próximo paso consiste en extender ese derecho a los efectos que tiene el hecho de ejercer los propios poderes. Aquí es donde la preocupación de Locke por el trabajo y el uso forman parte de la argumentación. Pero, como veremos, este segundo paso no da un derecho exclusivo a los frutos del propio trabajo, aun cuando sea el individuo quien aporta tales frutos.

Supongamos que en el estado de naturaleza yo cultivo una parcela de tierra con la intención de consumir lo que produzca. En ese caso, el ejercicio que yo haga de mis poderes es completamente independiente de cualquier otra persona y por lo tanto yo no mejoraría mi situación entrando en interacción con otros. Aun en el caso de que yo empeorara la situación de algún otro que de otro modo podría haber cultivado esa parcela, ese empeoramiento sería incidental al beneficio que yo recibo. Mi actividad no puede violar la estipulación. Ahora supongamos que alguna otra persona se apodera del producto de la tierra cultivado por mí. Esa persona mejora pues su situación como resultado de mi actividad y además empeora mi situación en comparación con lo que hubiera ocurrido en su ausencia, al privarme de los frutos de mi trabajo. Su actividad viola la estipulación. Por lo tanto vemos más claramente lo que decíamos en el ejemplo de 3.2; Jonathan no debe apoderarse de las frutas y verduras de Joanna y ésta no debe apoderarse del producto de la pesca de Jonathan.

Obsérvese que a fin de establecer una violación de la estipulación nos referimos tanto al trabajo que realicé para producir el bien que otra persona toma, como a mi intención de usar ese bien. La otra persona mejora su situación gracias a mi trabajo y empeora la mía al privarme del uso esperado. No se violaría la estipulación si esa persona tomara una fruta caída de un árbol que yo pensaba usar pero que no había hecho nada por producir ni tampoco si tomara algún fruto producido por mí pero que yo hubiera desechado.

Mediante este argumento demostramos un derecho sobre los efectos que tiene el trabajo de un individuo, pero no un derecho exclusivo a su posesión. Porque no mostramos que se violaría la estipulación si alguien tomara los frutos de mi trabajo y me compensara por mi esfuerzo y mi intención de usar tales frutos. Si el beneficio que recibo no es menor , desde el punto de vista de mis utilidades, del que yo podría haber esperado de mi trabajo en ausencia de ese comprador, mi situación no empeoraría.

Por consiguiente, la estipulación exige lo que, siguiendo a Nozick, llamamos la *compensación plena* y no la *compensación del mercado.*[14] La compensación plena hace que la persona no quede con ninguna pérdida neta de su utilidad. La compensación del mercado le proporciona una parte compartida del beneficio obtenido por el otro individuo que toma el bien y paga la compensación; la parte que cabría esperar del intercambio voluntario. Una persona empeora la situación de otra al no darle una compensación plena por los efectos de sus actos. Una persona puede mejorar la situación de otra dándole la compensación del mercado (en el caso en que ésta exceda la compensación plena); la ganancia que le brinda el intercambio es un beneficio que sólo puede recibir mediante la interacción. La estipulación prohíbe empeorar pero no exige mejorar la posición de otra persona, al mejorar la propia. Por lo tanto, no exige que la persona que toma los frutos del trabajo de otra comparta sus ganancias; sólo exige que la compense por los costos. Por consiguiente, la estipulación da un derecho *sobre* los frutos del propio trabajo y por lo tanto a la compensación plena, pero no da un derecho a aquellos frutos y por lo tanto tampoco a la compensación del mercado.

Este segundo paso, de un derecho al propio cuerpo y a sus poderes a un derecho sobre los propios productos, completa la conversión del estado puro de naturaleza. El paso restante tiene que ver con la transición de la interacción natural a la interacción del mercado y la interacción cooperativa. La primera de ellas constituye la internalización de los costos necesarios para que exista el mercado.

Supongamos que viviéramos como pescadores a orillas de un río. Si una persona me obliga a pescar para ella, viola la

estipulación porque me impide ejercer mis poderes como mejor me plazca. Si esa persona se apodera del pescado que yo capturo viola la estipulación a menos que me compense por mi trabajo y por mi intención de usar el pescado. Pero si la persona vive río arriba y sólo utiliza el río para echar en él sus desperdicios, aun cuando con esa actitud mate muchos peces de mi parte del río no viola la estipulación. Porque aunque empeore mi situación en comparación con lo que yo podría haber esperado en su ausencia, esa persona no mejora la suya en su interacción conmigo. No está mejor de lo que podría estar si río abajo no viviera nadie. El costo que esa persona me impone no necesariamente se relaciona con el beneficio que recibe, no es un *costo desplazado*. Antes bien, podemos decir que el costo se ocasiona únicamente por mi presencia, que, desde el punto de vista del otro, puede ser sencillamente una presencia no deseada.

Pero supongamos ahora que dejamos de vivir como pescadores independientes. En lugar de consumir todo el pescado que captura, mi vecino utiliza una parte para comerciar conmigo o implicándome a mí de algún modo. Mi voluntad de comerciar —mi deseo de comer pescado y los términos según los cuales acepto el pescado— sufren, por supuesto, la influencia de la oferta de pescado que tengo directamente disponible y por lo tanto de la actividad contaminante de mi vecino. El intercambio conmigo mejora su situación; es decir que, desde su punto de vista, la interacción conmigo es provechosa. Pero tal interacción no puede mejorar mi situación, si tomamos como punto de base la ausencia de mi vecino. Aunque yo me beneficie comerciando con él porque la alternativa es no comerciar, mi situación es peor de lo que sería si yo estuviera solo, pescando en una corriente sin desperdicios tóxicos. Teniendo en cuenta todas las formas posibles de interacción entre mi vecino y yo, él mejora su situación y yo empeoro la mía. Y por lo tanto, el uso que él hace del río como vaciadero viola la estipulación porque afecta los términos del comercio posible entre ambos.

Supongamos sin embargo que considerando nuestra interacción en su conjunto, cada uno de nosotros mejora su situación. Nuestros intercambios aparentemente me compensarían —plenamente— por el deterioro del río que produce el otro. Pero, ¿eso lo liberaría de la carga de violar la estipulación?

No; la ausencia de un empeoramiento global no muestra que alguna parte de nuestra interacción no viole la estipulación. Mi vecino desecha sus desperdicios de manera tal que mata peces que viven en mi parte del río. Por lo tanto, me impone un costo que le mejora los términos del comercio conmigo y en consecuencia los empeora para mí. El costo que me impone es ahora necesario para cierta parte del beneficio que recibe, por lo tanto es un costo desplazado. El otro se beneficia de arruinar mi agua; mejora su situación mediante una interacción que empeora la mía.

Nuestros intercambios no constituyen una compensación para mí por el agua contaminada. El vecino no comercia conmigo con el fin de compensarme. No comercia conmigo porque arruinó mi agua; su método de desembarazarse de sus residuos no es una condición necesaria para que ambos podamos comerciar. En realidad, nuestros intercambios, lejos de compensarme por la contaminación del agua, son menos beneficiosos para mí de lo que serían si no se hubiese arruinado el agua del río. Si mi vecino me habla de compensación agrega a la injuria el insulto. Me impone costos que no compensa mediante su forma de desembarazarse de sus residuos y, teniendo en cuenta nuestra interdependencia, se beneficia al imponerme esos costos; de modo que viola la estipulación.

Se viola la estipulación mediante una acción que mejora la situación del actor mediante el empeoramiento de la situación de otra persona. Una acción realizada en un contexto puede ser por completo inocente; puede beneficiar al actor e imponerle costos a otra persona, pero el beneficio y el costo son por completo independientes. La misma acción realizada en otro contexto puede ser claramente culpable; puede beneficiar al actor e imponerle costos a otra persona, y el beneficio, al menos en parte, puede depender del costo.

Al contaminar el agua de la corriente en la que pesco, mi vecino aumenta mi demanda por su pescado y así mejora, desde su punto de vista, los términos del comercio entre ambos. Pero no necesariamente ése es el único beneficio que obtiene a mis expensas. Supongamos que somos no ya simples pescadores, sino industriales. Ambos necesitamos agua pura para elaborar nuestros productos y ambos producimos residuos tóxicos de los

cuales la forma más barata de desembarazarse es echarlos al río y quitarle pureza al agua. Si cada uno de nosotros produce únicamente para su propio uso, el uso que hace mi vecino del río para desembarazarse de sus residuos aunque aumenta mis costos de producción (puesto que yo debo instalar un equipo purificador de agua en mi fábrica) no disminuye los suyos en comparación con una situación en la que yo estuviera ausente. Pero supongamos que ambos producimos para vender nuestros productos a una tercera persona. En ese caso, mi vecino se pone en situación de ventaja bajando los costos de desembarazarse de sus residuos de un modo que aumenta mis costos de producción. En esta situación de mercado, cualquier costo de la actividad del otro que recaiga en mí es un costo desplazado que lo beneficia a él y perjudica mi posición competitiva y así, al no estar compensado, constituye una violación a la estipulación.

Partiendo de este argumento, podemos generalizar y llegar a la conclusión de que tanto en la interacción de mercado como en la interacción cooperativa, todos los costos de las actividades de una persona que recaen en otras dentro de la esfera de la interacción son costos desplazados que exigen una compensación para que no se viole la estipulación. Porque aun cuando, estrechamente concebidos, algunos de esos costos puedan parecer innecesarios para los beneficios particulares recibidos, en la medida en que para esas actividades se necesite la estructura de las instituciones y las prácticas del mercado o las prácticas cooperativas, todos los costos que se den dentro de esa estructura son necesarios. Sin duda, dentro del mercado perfectamente competitivo o dentro de la interacción cooperativa regida por el principio de concesión relativa minimáxima no puede ocurrir ningún desplazamiento de los costos, pero para que ninguna de las partes saque ventaja inicial, sus posiciones iniciales deben ser el resultado de una interacción natural que satisfaga la estipulación. Por consiguiente, en la medida en que la interacción natural lleve a la interacción de mercado o a la interacción cooperativa, la estipulación exige la plena internalización de los costos entre aquellos que participan de la interacción. Aun cuando en el estado de naturaleza mismo, ciertos costos sólo se dan a causa de la presencia de otros que no afectan el beneficio recibido, en la medida en que estos otros formen parte de la

interacción social, los costos que se les imponen en el estado de naturaleza deben compensarse determinando sus dotes de mercado y sus dotes cooperativas.

Reiteramos este punto porque tiene gran importancia para nuestro análisis. En un estado puro de naturaleza, imponerles costos a aquellos que tienen una participación incidental en nuestras actividades, de ningún modo hace que violemos la estipulación. Tales costos son estrictamente incidentales e innecesarios respecto de los beneficios que recibimos. En un caso semejante nosotros no mejoramos nuestra situación mediante la interacción, aun cuando empeoremos la situación de aquellos sobre quienes recaen los costos. Pero si consideramos a esas personas como socios potenciales de las relaciones sociales que podemos establecer en la competencia del mercado y en la cooperación, la estipulación nos prohíbe pues imponerles ningún costo sin brindarles una compensación apropiada. Porque los costos que recaen sobre ellos los ponen en desventaja respecto de las relaciones futuras y por lo tanto mejoran nuestra situación y empeoran la de ellos.

Es necesaria la compensación plena y no la compensación de mercado. Lo que nos interesa son las condiciones previas del mercado y no la interacción misma del mercado. El pescador tenía un derecho sobre la corriente pero no —todavía— un derecho a ella. Tales derechos sólo aparecen cuando se da el paso final para convertir el estado de naturaleza en sociedad. El efecto de los tres primeros pasos destinados a racionalizar y moralizar el estado de naturaleza es la creación de una estructura de uso común entre las personas que participan de la interacción; una estructura en la cual es necesaria una compensación plena cuando una persona interfiere en el uso que otra persona hace o se propone hacer de ciertos bienes materiales, pero en la cual no hay derechos exclusivos de posesión de los objetos externos. El cuarto y último paso, que define las dotes plenas de cada individuo, introduce los derechos exclusivos a la tierra y a otros bienes.

Supongamos que varias personas habitan en una isla. La tierra y sus recursos constituyen en efecto los pastos comunes disponibles para todos. Pero el uso es individual; cada persona se ocupa primariamente de satisfacer sus propias necesidades

y la interacción no es cooperativa. A fin de que nuestro ejemplo sea más realista vamos a considerar que los habitantes son no individuos sino familias. La idea de una familia es la idea de un grupo de miembros que se interesan mutuamente unos por los otros; de ahí que la interacción interna dentro de la familia no sea relevante para nuestro análisis. Suponemos pues que cada familia ha de ocuparse de satisfacer sus propias necesidades, manteniendo una interacción no cooperativa con las demás familias. Pero un individuo o jefe de una familia, sabiendo que el cultivo planificado e intensivo ha de hacer más productiva la tierra, propone que se le asigne una porción de la isla para su uso exclusivo, de modo tal que él (y su familia) pueda beneficiarse maximizando su productividad. Intenta obtener un derecho exclusivo a cierta porción de la isla.

¿Cómo podemos evaluar la proposición de este habitante de la isla? Primero debemos preguntarnos si alguien, al pretender el uso exclusivo de la tierra y sus bienes, viola la estipulación, mejorando su situación a costa del empeoramiento de la situación de los demás. Si la respuesta es no, debemos preguntarnos si alguna otra persona, al interferir con esa pretensión al uso exclusivo, viola la estipulación. Si la respuesta es sí, el derecho propuesto queda establecido. El obtuvo el derecho y ningún otro puede utilizar la tierra o los bienes sin violar la estipulación. Y si se viola la estipulación, la compensación exigida ha de ser no la plena, sino la de mercado, porque el que obtuvo el derecho lo logró como podría haberlo logrado mediante el intercambio voluntario, como compensación por cualquier uso hecho de su tierra o sus bienes. El derecho a la tierra o los bienes es un derecho no sólo a los frutos de uso, sino también a los frutos de intercambio.

Comencemos pues por considerar los efectos de aceptar una pretensión de control exclusivo. Por un lado es evidente que la persona, a la que llamaré Eva, intenta mejorar su situación apropiándose de la tierra. E intenta mejorarla en relación con el punto base establecido por los términos del problema, es decir, en relación con el sistema de uso común. Eva busca la seguridad de la tenencia, que le confiere un derecho a la compensación de mercado, antes que a la compensación meramente plena. Por lo tanto intenta mejorar su situación respecto de sus semejantes.

Y esta conclusión se refuerza al observar que un sistema de derechos determina las dotes para las futuras interacciones de mercado y cooperativa; tener derechos exclusivos a cierta porción de la tierra le proporciona pues a Eva bases más favorables para tal interacción que las que podría tener de otro modo.

Pero aunque Eva intente mejorar su situación respecto de sus semejantes, no necesita hacerlo empeorando la situación de aquéllos. Es verdad que ellos han de perder su derecho al uso en común de la tierra de la que se apodera Eva, y si desean usar lo que ella produzca han de verse obligados a mantener los intercambios que Eva acepte voluntariamente antes que pagarle meramente su compensación plena. Ahora bien, podemos suponer que Eva pretenda una parte de la isla tan amplia que deje a sus compañeros peor que antes; que la tierra que les quede en común no los mantenga tan bien como los mantenía la isla entera. En ese caso la pretensión de Eva violaría la estipulación. El cultivo planificado e intensivo que pudo lograrse mediante la seguridad de la tenencia, bien puede permitirle vivir mejor a Eva en una parte de la isla suficientemente pequeña para que los demás también puedan vivir tan bien sin Eva en la tierra restante, como vivían cuando todos usaban la isla entera en común. Porque, por supuesto, al pretender una posición privada, Eva propone perder su derechos a los pastos comunes restantes.

Además, los otros habitantes de la isla también pueden beneficiarse de las nuevas oportunidades de comerciar sus productos por algunos de los bienes resultantes del cultivo más intensivo de Eva. Eva puede producir los alimentos suficientes para satisfacer las necesidades de varias familias, de modo tal que los otros, que antes cultivaban su propio alimento, puedan convertirse en expertos artesanos, con beneficio para todos. De ahí que el hecho de que Eva se apropie de esa porción de tierra puede mejorar la situación de todos respecto del punto de base establecido por el uso en común; entonces está claro que la actitud de Eva no viola la estipulación.

Queda pues defendido el derecho de Eva. Los derechos exclusivos de posesión pueden aportar beneficios a todos, porque les dan a los individuos la seguridad necesaria para que les sea provechoso usar los recursos que tienen disponibles los seres

y la interacción no es cooperativa. A fin de que nuestro ejemplo sea más realista vamos a considerar que los habitantes son no individuos sino familias. La idea de una familia es la idea de un grupo de miembros que se interesan mutuamente unos por los otros; de ahí que la interacción interna dentro de la familia no sea relevante para nuestro análisis. Suponemos pues que cada familia ha de ocuparse de satisfacer sus propias necesidades, manteniendo una interacción no cooperativa con las demás familias. Pero un individuo o jefe de una familia, sabiendo que el cultivo planificado e intensivo ha de hacer más productiva la tierra, propone que se le asigne una porción de la isla para su uso exclusivo, de modo tal que él (y su familia) pueda beneficiarse maximizando su productividad. Intenta obtener un derecho exclusivo a cierta porción de la isla.

¿Cómo podemos evaluar la proposición de este habitante de la isla? Primero debemos preguntarnos si alguien, al pretender el uso exclusivo de la tierra y sus bienes, viola la estipulación, mejorando su situación a costa del empeoramiento de la situación de los demás. Si la respuesta es no, debemos preguntarnos si alguna otra persona, al interferir con esa pretensión al uso exclusivo, viola la estipulación. Si la respuesta es sí, el derecho propuesto queda establecido. El obtuvo el derecho y ningún otro puede utilizar la tierra o los bienes sin violar la estipulación. Y si se viola la estipulación, la compensación exigida ha de ser no la plena, sino la de mercado, porque el que obtuvo el derecho lo logró como podría haberlo logrado mediante el intercambio voluntario, como compensación por cualquier uso hecho de su tierra o sus bienes. El derecho a la tierra o los bienes es un derecho no sólo a los frutos de uso, sino también a los frutos de intercambio.

Comencemos pues por considerar los efectos de aceptar una pretensión de control exclusivo. Por un lado es evidente que la persona, a la que llamaré Eva, intenta mejorar su situación apropiándose de la tierra. E intenta mejorarla en relación con el punto base establecido por los términos del problema, es decir, en relación con el sistema de uso común. Eva busca la seguridad de la tenencia, que le confiere un derecho a la compensación de mercado, antes que a la compensación meramente plena. Por lo tanto intenta mejorar su situación respecto de sus semejantes.

Y esta conclusión se refuerza al observar que un sistema de derechos determina las dotes para las futuras interacciones de mercado y cooperativa; tener derechos exclusivos a cierta porción de la tierra le proporciona pues a Eva bases más favorables para tal interacción que las que podría tener de otro modo.

Pero aunque Eva intente mejorar su situación respecto de sus semejantes, no necesita hacerlo empeorando la situación de aquéllos. Es verdad que ellos han de perder su derecho al uso en común de la tierra de la que se apodera Eva, y si desean usar lo que ella produzca han de verse obligados a mantener los intercambios que Eva acepte voluntariamente antes que pagarle meramente su compensación plena. Ahora bien, podemos suponer que Eva pretenda una parte de la isla tan amplia que deje a sus compañeros peor que antes; que la tierra que les quede en común no los mantenga tan bien como los mantenía la isla entera. En ese caso la pretensión de Eva violaría la estipulación. El cultivo planificado e intensivo que pudo lograrse mediante la seguridad de la tenencia, bien puede permitirle vivir mejor a Eva en una parte de la isla suficientemente pequeña para que los demás también puedan vivir tan bien sin Eva en la tierra restante, como vivían cuando todos usaban la isla entera en común. Porque, por supuesto, al pretender una posición privada, Eva propone perder su derechos a los pastos comunes restantes.

Además, los otros habitantes de la isla también pueden beneficiarse de las nuevas oportunidades de comerciar sus productos por algunos de los bienes resultantes del cultivo más intensivo de Eva. Eva puede producir los alimentos suficientes para satisfacer las necesidades de varias familias, de modo tal que los otros, que antes cultivaban su propio alimento, puedan convertirse en expertos artesanos, con beneficio para todos. De ahí que el hecho de que Eva se apropie de esa porción de tierra puede mejorar la situación de todos respecto del punto de base establecido por el uso en común; entonces está claro que la actitud de Eva no viola la estipulación.

Queda pues defendido el derecho de Eva. Los derechos exclusivos de posesión pueden aportar beneficios a todos, porque les dan a los individuos la seguridad necesaria para que les sea provechoso usar los recursos que tienen disponibles los seres

humanos de maneras más eficaces y productivas. Los derechos exclusivos transforman un sistema en el cual cada uno trabaja en los pastos comunes a fin de satisfacer sus propias necesidades en un sistema en cual cada uno trabaja en su propiedad y satisface sus necesidades mediante el intercambio de mercado. La autosuficiencia individual da lugar a la especialización de las funciones. La división del trabajo abre nuevos estilos de vida, con oportunidades y satisfacciones anteriormente ni siquiera imaginadas. Por lo tanto la naturaleza mutuamente beneficiosa de los derechos exclusivos de la posesión proporciona una base suficiente para que tales derechos surjan de la condición de uso común, que es la forma final del estado de naturaleza. Esos derechos dependen de la estipulación que permite la apropiación individual y prohíbe la posterior interferencia. Cada persona está plenamente individualizada y definida en relación con sus semejantes y su derecho al propio cuerpo y a sus poderes se extiende ahora a los bienes materiales. Cuando el mercado fracasa están disponibles las formas de cooperación que permiten alcanzar la optimidad ante las externalidades. Eva, la primera que se apropió de un bien común, es la gran benefactora de la humanidad, aunque haya sido impulsada por una mano invisible a promover un fin que no formaba parte de sus intenciones.

Por supuesto las diferentes personas han de beneficiarse de manera diferente de la aparición de un sistema de derechos exclusivos. Podemos suponer que Eva, que fue la primera en tomar la tierra para su uso exclusivo, ha de quedarse con la mejor porción; ninguna otra persona es pues capaz de apropiarse de una porción igualmente ventajosa. Eva no deja a sus semejantes algo "tan bueno" para apropiarse, aunque al tomar su porción para sí deje a los demás tan bien, o mejor, que antes. La estipulación asegura que en cada etapa de la interacción, cada persona quede por lo menos tan bien como podía esperar estar desde su posición anterior. Por lo tanto, nadie saca ventaja, pero tampoco queda asegurada la equidad. Debemos mostrar pues que la inequidad que permite la estipulación no es una indicación de parcialidad.

4.2 ¿Constituye la estipulación una restricción a la interacción? ¿Qué podría llevarnos a cuestionar su imparcialidad? ¿Va demasiado lejos al prohibir a las personas que mejoren su situación mediante una interacción que empeore la situación de otros? ¿No va lo suficientemente lejos al dejar a las personas en libertad de mejorar su situación siempre que no obtengan ganancias del empeoramiento de algún otro? ¿O al dejar que las personas empeoren libremente la situación de los demás si la única alternativa fuera empeorar la propia? ¿O la estipulación carece de imparcialidad porque se basa en consideraciones referentes al mejoramiento y el empeoramiento? Esta última pregunta pone en el tapete serias cuestiones a las que debemos referirnos. Al cuestionar la imparcialidad de la estipulación lo que se está cuestionando es la importancia de mejorar y empeorar las situaciones de los actores.

Porque —podría sostenerse— la estipulación no dice nada respecto de la equidad. O, también podría decirse, la estipulación no dice nada respecto de satisfacer necesidades. El rico puede comer opíparamente caviar y beber champagne mientras una mujer pobre muere de hambre a su puerta. Y esa mujer ni siquiera puede tomar las migas de la mesa del rico si con ello priva a éste del placer de alimentar a sus pájaros.

Por penosa que pueda parecernos esta situación, no debemos dejarnos engañar. Pensamos en ricos y pobres dentro de un contexto social y pensamos que la riqueza de ese hombre y la pobreza de esa mujer están de algún modo relacionados. Y si así fuera, al examinar cómo se llegó a esa situación, podríamos muy bien comprobar una violación, si no ya de la estipulación, del principio de concesión relativa minimáxima. Deberíamos comenzar con un ejemplo por completo diferente, un ejemplo que tomaremos de *Anarchy, State, and Utopia* de Robert Nozick y que adaptaremos a nuestro objetivo.[15]

Hay dieciséis Robinson Crusoe que viven en dieciséis islas diferentes. Cada uno de ellos es o bien inteligente, o bien tonto, o bien fuerte o bien débil, o bien enérgico o bien holgazán. Cada uno de ellos vive o en una isla bien provista o en una mal provista de lo necesario para satisfacer las necesidades y deseos humanos. Ninguno de nuestros Robinson Crusoe es idéntico a otro ni en sus características ni en sus circunstancias (al haber dieciséis de ellos se dan todas las combinaciones posibles). Hay un Crusoe inteligente, fuerte y enérgico que vive muy cómodamente en

una isla bien provista. Hay otro tonto, débil y perezoso que apenas sobrevive en una isla mal provista. Y entre ellos hay otros catorce con distintas características y en diferentes situaciones.

Cada uno de ellos cuenta con un aparato de radio emisor receptor que le permite comunicarse con los demás Robinson Crusoe. Cada uno sabe cuál es su situación en comparación con la de los demás. Cada uno puede construir pequeñas balsas, lo suficientemente amplias para transportar provisiones de diferentes tipos, pero no para llevar a un hombre o una mujer. Las corrientes oceánicas impulsan esas balsas de una isla a otra, pero en una sola dirección, de modo que es imposible comerciar.

Supongamos ahora que el Crusoe inteligente, fuerte y enérgico que vive en la isla bien provista pueda enviarle provisiones en su balsa al Crusoe tonto, débil y perezoso que vive en la isla mal provista. Sin duda sería generoso de su parte hacerlo. Y quizá quisiera hacerlo; en nuestro relato nada indica que el primer Crusoe no se interese por los intereses de los demás. Pero supongamos que no desea hacerlo; sabe cuál es la situación de los demás pero no se preocupa por ellos. ¿Tiene acaso la obligación de darle provisiones al Crusoe tonto, débil y perezoso? ¿Sería injusto si no lo hace? El principio que establece que cada Crusoe puede utilizar sus propias habilidades y los recursos de su isla en su propio provecho, ¿es un principio injusto o parcial? ¿Un principio imparcial exigiría a los Crusoe mejor situados contribuir a mejorar la situación de los peor dotados, en el caso de que las corrientes oceánicas lo permitieran? ¿Exigiría contribuciones equitativas si fuera posible hacerlas? ¿O establecería las necesidades básicas y exigiría que en la medida de lo posible éstas quedaran satisfechas? No. Ningún principio salvo el que establece que cada Robinson Crusoe puede buscar su propio provecho sería justo e imparcial, porque cualquier otro permitiría la aparición de jugadores líberos o de parásitos.

Transcurre el tiempo y los Crusoe aprenden a construir balsas más amplias con lo cual pueden emigrar (en la dirección de las corrientes). Si nos atenemos a la estipulación, cada uno de ellos tiene un derecho a sus facultades físicas y mentales, pero sólo un derecho sobre los recursos de la isla en la que habita.

(Teniendo en cuenta la diferencia entre las islas bien provistas y las mal provistas, una vez que se puede emigrar, algunos Crusoe podrían mejorar su situación adquiriendo un derecho exclusivo sobre una isla de un modo que podría empeorar la situación de algún otro.) De ahí que al poder emigrar y poder utilizar los recursos de cualquier isla en la que atraque, el Crusoe que emigra debe compensar (plenamente) a cualquier otro al que le quite el fruto de su trabajo. El Crusoe inteligente, fuerte y enérgico que vive en la isla mal provista (si la corriente lo permite) puede trasladarse a la isla bien provista habitada por el Crusoe tonto, débil y holgazán e, ignorando a este último, instalarse a vivir tan cómodamente como el otro Crusoe inteligente, fuerte y enérgico que comenzó su vida en la isla bien provista. Por supuesto, algunas migraciones podrían conducir a la cooperación; por ejemplo, el Crusoe inteligente, débil y perezoso podría combinar sus habilidades con las del Crusoe fuerte y enérgico pero tonto. Pero no necesariamente tiene que ocurrir esto. ¿Es injusto el principio que permite que cada Crusoe utilice sus propias habilidades y los recursos de cualquier isla a la que pueda llegar en su propio provecho? ¿Podría un principio imparcial exigir que los Crusoe más capaces se trasladen (si las corrientes lo permiten) para ayudar a los menos hábiles? No. Semejante exigencia sería injusta, pues una vez más permitiría la aparición de jugadores líberos y de parásitos.

John Rawls afirma: "No hay más razón para admitir que los ingresos y la riqueza se distribuyan según la distribución natural de los haberes que según la fortuna histórica y social".[16] Antes bien "deberíamos considerar la distribución de los talentos naturales como un haber común y...compartir los beneficios de esa distribución, sea lo que fuere que resulten ser. Aquellos que fueron favorecidos por la naturaleza, sean quienes fueren, pueden obtener una ganancia de su buena fortuna sólo sobre bases que mejoren la situación de los que resultaron menos favorecidos... Nadie merece esa mayor capacidad natural ni un punto de partida más favorable en la sociedad".[17] Está claro que para Rawls las personas deberían negociar partiendo de una posición de igualdad, no sólo respecto de los derechos iniciales a los bienes, sino también respecto de los derechos iniciales a los poderes y habilidades personales. Aquellos más capaces o más

afortunados no sólo deben frenar el impulso de sacar ventaja de los demás como lo exige la estipulación, sino que además deben darles ventaja si pretenden beneficiarse. Para Rawls la moral exige que se permita la existencia de los jugadores líberos; no podemos interpretar de otro modo su idea de que se considere que los talentos personales constituyen un haber común.

Podemos coincidir con Rawls en lo referente a que nadie merece su capacidad natural. Siendo cada persona como es, no se trata de una cuestión de merecimiento. Pero, ¿qué se sigue de esto? En una condición de soledad, la capacidad natural de una persona determina lo que esa persona obtiene, dadas sus circunstancias. La propia capacidad natural es lo que cada individuo aporta a la sociedad, a la interacción de mercado y cooperativa. ¿Por qué razón no debería determinar, o ayudar a determinar, lo que cada individuo recibe de la sociedad? ¿Cómo puede un principio determinar imparcialmente la manera en que las personas han de beneficiarse en la interacción, excepto teniendo en cuenta lo que esas personas hubiesen obtenido si no hubiese habido interacción?

Rawls habla de la "distribución de los talentos naturales" y de la "lotería" natural.[18] Y al hacerlo cae víctima de los peligros que acechan detrás de esas expresiones. No existe una lotería natural; nuestros talentos no se distribuyen partiendo de una cantidad fija que asegure la existencia de ganadores y de perdedores. Y si bien hay una distribución, no hay un distribuidor —a menos que aceptemos una base teísta, completamente extraña al argumento de Rawls—. La estipulación determina las dotes iniciales de las personas que participan en la interacción y tiene en cuenta las diferencias reales que existen entre esas personas consideradas como actores. Pero la estipulación no es en sí misma un actor o un distribuidor; se incorpora a las acciones de las personas en la medida en que éstas la acaten. Si hubiese un distribuidor de los haberes naturales o si la distribución de los recursos disponibles fuera el resultado de una elección social, podríamos pues suponer razonablemente que en la medida de lo posible, las porciones repartibles deberían ser iguales y que sólo se justificaría una distribución desigual cuando redundara en beneficio de todos. Pero esto equivaldría a considerar a las personas como las criaturas de un distribuidor

—un Dios o una Sociedad no instrumental— y no como actores racionales e individuales. Si bien estamos de acuerdo con Rawls en cuanto a que la sociedad es una empresa cooperativa destinada a la ventaja mutua, debemos disentir en cuanto a que los talentos naturales deban considerarse haberes comunes. Los dos enfoques ofrecen concepciones antitéticas tanto del ser humano individual como de la sociedad.

Cada ser humano es un actor con ciertas preferencias y ciertas aptitudes físicas y mentales, facultades que, en ausencia de sus semejantes, utiliza directamente para satisfacer tales preferencias. Esto proporciona una base, en absoluto arbitraria, partiendo de la cual podemos examinar y evaluar la interacción introduciendo conceptos tales como el mejoramiento y el empeoramiento. Un principio extraído de esta base no relacionarían a los seres humanos entre sí como actores. Un principio que no tomara esa base como fundamento normativo no relacionaría imparcialmente a los seres humanos como actores.

Si los dieciséis Robinson Crusoe vivieran, cada uno en una isla, y utilizaran sus propias capacidades para apropiarse de los recursos de la isla, el resultado, sea el que fuere, estaría justificado. También estaría justificado si cada persona aportara, como dote individual, sus capacidades y todo aquello que haya conseguido mediante tales habilidades sin empeorar la situación de los demás, a los esfuerzos competitivos y cooperativos que constituyen la sociedad. Al determinar los derechos que tienen las personas sobre la base de lo que tales personas hacen, y al considerar lo que las personas hacen desde el punto de vista del actor individual, la estipulación asegura la imparcialidad de la interacción.

4.3 Los seres humanos tienen derechos. La teoría moral no ha tenido un éxito muy notable al tratar de dar una explicación de tales derechos. Como vimos en 4.4.1, el utilitarismo lleva inexorablemente a pensar que "los derechos... se depositan en el gobierno y la asociación de éste con los individuos es una cuestión de conveniencia temporal".[19] Si este enfoque no fuera inaceptable, deberíamos preguntarnos cómo llegaron los gobiernos a adquirir o poseer los derechos, una pregunta bastante difícil de responder si uno supone que los gobiernos son instrumentos antes que amos de sus ciudadanos, pero podemos correr

un piadoso manto de oscuridad sobre esta *gaffe* de los utilitaristas. Los defensores a ultranza de los derechos individuales, a su debido tiempo, se contentaron con dejar de lado la tarea de justificar los derechos, "siguiendo la respetable tradición de Locke que nunca proporcionó nada ni remotamente parecido a una explicación satisfactoria ni de la condición ni de los fundamentos de la ley natural".[20] En realidad Locke, a diferencia de sus discípulos actuales, no suponía que fuera posible dar una fundamentación secular de los derechos individuales; su teoría moral, a diferencia de la de Hobbes, es abiertamente teísta.

Los defensores del contrato ofrecen una explicación secular de los derechos. Pero puede malinterpretarse la idea de una moral basada en el acuerdo si se supone que los derechos deben ser el producto o el resultado del acuerdo. Si aceptáramos ese enfoque, deberíamos suponer que el principio de concesión relativa minimáxima determina los derechos. Pero, como ya vimos, para que pueda aplicarse este principio, o de manera más general, para que puedan darse la interacción cooperativa o la interacción de mercado, es necesario hacer una definición inicial de los actores, atendiendo a sus recursos disponibles, y nosotros identificamos los derechos individuales con esos recursos. Los derechos proporcionan el punto de partida necesario para iniciar el acuerdo y no el resultado de éste. Los derechos constituyen lo que cada persona aporta a la mesa de negociación y no lo que recibe de ella.

Las prácticas del mercado y las prácticas cooperativas suponen la existencia de derechos individuales. Esos derechos quedan moralmente establecidos en la estipulación. Y los derechos así fundamentados resultan ser los derechos familiares de nuestra tradición: los derechos a la persona y a la propiedad. En la sección 5 mostraremos que además están fundamentados desde el punto de vista racional.

Sin embargo, debemos reconocer que tales derechos no son inherentes a la naturaleza humana. Al definir a las personas como partícipes de la competencia de mercado y de la cooperación, esos derechos asignan la prioridad moral del individuo a la sociedad y sus instituciones. Pero no le proporcionan a cada individuo una condición moral inherente en relación con sus semejantes. En un estado de naturaleza puro, en el cual las

personas se interrelacionan de manera no cooperativa y sin perspectivas de cooperación, tales derechos no tienen cabida. Rawls dice que "la sociedad es una empresa cooperativa destinada a lograr la ventaja mutua".[21] Y esa perspectiva de ventaja mutua es lo único que hace que los derechos entren en el juego, como restricciones a la conducta individual. Esa perspectiva permite que los derechos coexistan con el supuesto del desinterés mutuo. Las pretensiones morales que cada uno de nosotros le presenta a los otros y que están expresadas en nuestros derechos, no dependen ni de nuestros afectos recíprocos ni de nuestras capacidades racionales o intencionales, como si éstas exigieran un acatamiento inherente, sino que dependen de nuestra participación real o potencial en actividades que producen beneficio mutuo.

Al tratar de maximizar su utilidad sujetándose a dos niveles de restricción, cada persona obra racionalmente. Primero, como lo determina la estipulación, cada individuo está restringido por los derechos de sus semejantes. Si cada uno respeta los derechos de los demás, nadie saca pues ventaja de ningún otro. Quedan satisfechas las condiciones iniciales para que se dé la competencia de mercado justa y racional y para poder participar de la negociación —explícita o implícita— que establece los términos de una cooperación racional y justa. Segundo, cada individuo está restringido por las exigencias de la concesión relativa minimáxima, dentro de las instituciones y prácticas cooperativas. El error que cometen algunos defensores de los derechos individuales es suponer que los derechos, el primer nivel de restricción, bastan por sí solos y, en realidad, suponer que deben rechazarse otras restricciones morales adicionales a la maximización de la utilidad individual por ser incompatibles con tales derechos. Esos pensadores estarían en lo cierto si las personas que se adhieren a la estipulación participaran siempre de la interacción con sus semejantes en condiciones de competencia perfecta. Porque, como mostramos en el capítulo 4, el mercado es una zona moralmente libre. Pero las externalidades que tiñen nuestra interacción nos obligan a reconocer que los derechos, aunque necesarios, son insuficientes, desde el punto de vista moral. El tema básico de la idea de una moral fundada en el acuerdo es demostrar que los dos

niveles de restricción, el de los derechos y el de la adhesión a las prácticas cooperativas, son compatibles.

5.1 ¿Por qué un individuo que trata de maximizar su utilidad debería refrenar su impulso de sacar ventaja de sus semejantes? ¿Por qué su interés en los arreglos cooperativos y de mercado (que es en sí mismo una expresión de su interés por maximizar la utilidad) constituye una razón para aceptar la estipulación, como una forma de limitar sus derechos en la interacción cooperativa y de mercado? A fin de ilustrar mejor estas preguntas, volvamos a considerar el caso de nuestros pescadores. El que vive río arriba arroja sus residuos tóxicos a la corriente de agua, con lo cual la contamina y aumenta los costos que debo sufrir yo que vivo río abajo. Este hecho lo beneficia en su interacción conmigo y por lo tanto hace entrar en juego la estipulación; mi vecino no tiene derecho a contaminar el agua.

Debemos considerar dos alternativas: o bien el uso que le da mi vecino al río como vaciadero es el método más eficiente para todos o bien no lo es. Si realmente lo es, esa persona debería continuar contaminando el agua pero debería compensarme por los costos que me provocaría. Si no es el método más eficiente para todos, mi vecino debería adoptar otro método de desembarazarse de los residuos tóxicos que minimice los costos totales y pagarlo. En ambos casos se logra la optimidad. Pero puede obtenerse un resultado óptimo de un modo por completo diferente. Si utilizar el río para desechar sus residuos es eficiente, mi vecino debería sencillamente continuar haciéndolo. Si no lo es, yo debería pagarle la diferencia de los costos necesarios para inducirlo a adoptar un método más eficiente, puesto que ese precio ha de ser menor para mí que el costo de la contaminación. Así también se llega a un resultado óptimo, aunque por supuesto los costos y los beneficios quedan distribuidos entre nosotros de manera muy diferente.

Si mi vecino y yo acordaramos conformar una sociedad, él sin duda querría que continuaran las prácticas existentes y yo sostendría que quien produce los residuos debe hacerse cargo de los costos de desembarazarse de ellos. Mi vecino consideraría que la contaminación que produce es una simple consecuencia de la interacción natural, y no consideraría una condición necesaria para la cooperación el hecho de frenar su actividad

contaminante; aun cuando la frenara si yo lo indujera a hacerlo. Yo consideraría que los efectos de la contaminación constituyen un costo desplazado hacia mí en la interacción natural y no tendría razón para aceptarla como una condición de la cooperación; aun cuando la aceptara a cambio de una compensación. ¿Cómo podemos conciliar estos puntos de vista opuestos? Puesto que estamos intentando encontrar la racionalidad de la estipulación, podríamos apelar a ella para zanjar la cuestión. La estipulación le niega a mi vecino el derecho a contaminar las aguas y exige que se me compense si de todos modos la contaminación ocurre. Pero, ¿acaso las características de esta situación no indican una asignación muy diferente de los derechos? Mi vecino no ejerce una coacción sobre mí al desembarazarse de sus residuos mediante ese método; no me asigna una transferencia improductiva mediante medios depredadores. Sencillamente lanza sus residuos en el río. Si yo tengo alguna objeción que hacer, ¿no me corresponde a mí acaso hacer algo al respecto? Si la prohibición de la coerción beneficia a alguien, parece favorecer a mi vecino. De ahí que, como un maximizador racional de la utilidad, él no tenga razones para cambiar su método de deshacerse de sus residuos tóxicos, y ciertamente tampoco tenga razones para hacerme a mí un pago transferido, improductivo en sí mismo, como "compensación", a menos que yo lo obligue a pagar. Parecería que la razón y la estipulación se separan y así, puesto que la estipulación garantiza la imparcialidad, también lo hacen la razón y la moral.

Pero nosotros negamos que sea así. Comencemos nuestra réplica recordando que la cooperación tiene como única y suficiente justificación racional la maximización de la utilidad esperada. Por consiguiente, en la negociación la pretensión presentada y la concesión ofrecida por cada individuo dependen de su esfuerzo por maximizar su utilidad, junto con el reconocimiento de los esfuerzos similares que hacen los demás. El principio de concesión relativa minimáxima determina el resultado de la interacción cooperativa, de modo tal que las porciones del excedente cooperativo que han de compartirse están relacionadas con las contribuciones hechas para su producción, del mismo modo para todos. Por supuesto, no toda interacción particular, considerada independientemente de una práctica de

cooperación, ha de beneficiar a cada parte de manera proporcional a su contribución. Todos los miembros pueden esperar obtener una ganancia de ciertos arreglos destinados a la asistencia mutua, aun cuando en determinadas ocasiones el que recibe la asistencia gane y el que la brinda pierda. No es la ocasión sino la práctica lo que debe satisfacer la concesión relativa minimáxima.

Volvamos a los pescadores. Yo sufro una pérdida neta si mi vecino arroja sus residuos tóxicos al río. Desembarazarse de los desperdicios mediante el método que le resulte menos costoso a quien lo produce, ignorando los efectos que pueda tener su conducta en los demás, no es una práctica que ofrezca una beneficio esperado a cada miembro de la sociedad. No puede pues defenderse la interacción particular relacionándola con una práctica que satisfaga la concesión relativa minimáxima. De ahí que tal práctica viole la exigencia, fundamental para que exista una cooperación racional, del beneficio mutuo proporcional a la contribución.

Si la interacción está destinada a ser plenamente cooperativa, debe proceder de una posición inicial en la cual los costos estén internalizados y por lo tanto, de una posición en la cual ninguna persona tenga el derecho de imponerle a otra costos no compensados. Porque si no es así, los acuerdos sociales resultantes han de producir interacciones imparciales que beneficien a algunas personas a expensas de otras. Aun cuando cada miembro fuera a recibir cierta porción del excedente cooperativo, cada individuo no podría esperar beneficiarse en la misma proporción correspondiente a la contribución hecha que los demás. Las interacciones basadas en costos desplazados serían redistributivas y la redistribución no puede formar parte de un sistema de cooperación racional.

Pero, ¿por qué los maximizadores de la utilidad racionales deberían participar de manera plenamente cooperativa en la interacción? O de manera más general, ¿por qué deberían los individuos racionales participar plenamente de la sociedad, el lugar donde se desarrollan tanto la interacción de mercado como la interacción cooperativa, en vez de aceptar las prácticas particulares cooperativas y de mercado que se darían en un estado de naturaleza permanente? En efecto, ¿no es acaso el

estado de naturaleza la realidad que subyace en la condición humana? Estas preguntas nos remiten a la posición del Necio. Y aquí, como antes, nuestra respuesta al Necio se limita a enumerar las condiciones necesarias para que el acatamiento de las prácticas cooperativas sea racional. En el capítulo 6 el Necio, al referirse a la maximización directa propia del estado de naturaleza, negaba la racionalidad de acatar las estrategias conjuntas acordadas. Y nosotros mostramos que un individuo racional se dispone a tal acatamiento. Aquí el Necio niega la racionalidad de limitar los efectos que produce la interacción natural en el hecho de determinar los derechos individuales de los cuales procede la interacción social. Ahora debemos mostrar que, sin limitaciones que excluyan la posibilidad de sacar ventaja, un individuo racional no podría disponerse a abrazar la interacción cooperativa.

En 6.2.4 hicimos una distinción entre el acatamiento amplio y el acatamiento estrecho. Una persona dispuesta a acatar ampliamente compara el beneficio que podría esperar de la cooperación (sean cuales fueren los términos ofrecidos) con lo que podría esperar de la no cooperación, y acata únicamente si el primer beneficio es mayor que el segundo. Si todas las personas estuviesen dispuestas a acatar ampliamente, ninguna tendría razones para aceptar la estipulación, ni ninguna otra restricción a la interacción natural. El resultado no cooperativo serviría de punto de base y cualquier mejora que pudiera ofrecerse sobre ese punto atraería la cooperación voluntaria. Pero el acatamiento amplio no es una disposición racional para los maximizadores de la utilidad. Una persona que acata ampliamente invita a que los demás le saquen ventaja al establecer los términos de la cooperación, y si hay varias personas que acatan ampliamente, los demás al entrar en interacción con ellas comprobarían que es conveniente y hasta más ventajoso no acatar ampliamente. Si alguien está dispuesto a acatar a fin de obtener el beneficio que sea, al entrar en interacción con él yo debería disponerme a aceptar una estrategia conjunta sólo si esa estrategia me ofrece no una porción justa, sino la parte del león del superávit cooperativo. Por lo tanto no es ni puede ser racional para nadie disponerse a abrazar el acatamiento amplio. Pero, puesto que nadie elige

restringir su conducta por propia iniciativa, nadie considera racional acatar los acuerdos más que los otros. Una igual racionalidad exige un acatamiento igual. Puesto que el acatamiento amplio no es racional para todos, tampoco lo es para cada individuo.

Una persona dispuesta a abrazar el acatamiento estrecho compara el beneficio que podría esperar de la cooperación con el que podría obtener de un resultado justo y óptimo, y acata una estrategia conjunta sólo si el primer beneficio se aproxima al segundo. Un resultado es justo si satisface las normas de imparcialidad, las cuales, como ya lo mostramos, quedan establecidas por la estipulación y el principio de concesión relativa minimáxima. Por lo tanto, una persona dispuesta al acatamiento estrecho espera que los demás acepten y consideren racional aceptar la estipulación como una condición de la cooperación. Pero entonces, atendiendo a que todos tienen la misma racionalidad, también ella debe considerar racional aceptarla como una condición de la cooperación. Por consiguiente, la disposición al acatamiento estrecho incluye la disposición a aceptar la estipulación como una restricción a la interacción natural, en la medida en que el individuo desee formar parte de la sociedad y participar de las prácticas cooperativas y de mercado.

La racionalidad de disponerse a acatar estrechamente y por lo tanto a aceptar la estipulación, procede del carácter ventajoso que tiene la sociedad y de que sus miembros gocen de igual racionalidad. Si todas las personas no estuvieran siquiera dispuestas al acatamiento estrecho y se negaran a adherise voluntariamente a las estrategias conjuntas que produzcan resultados justos y óptimos, la cooperación no sería pues posible. Si algunas personas no llegaran a acatar ni siquiera estrechamente, la cooperación sería pues posible sólo si los otros acataran algo más. Pero esto violaría la igual racionalidad. Si algunas personas acataran algo más que estrechamente, los demás considerarían ventajoso, y por lo tanto racional, acatar menos. Pero esto también viola la igual racionalidad. Para cada individuo es racional acatar los acuerdos lo suficiente para que la sociedad sea posible, sólo si los demás hacen lo mismo que él. Para nadie es racional acatar los acuerdos lo suficiente para que

la sociedad sea posible si los demás son menos complacientes. Por consiguiente, para cada individuo es racional acatar los acuerdos estrechamente.

Cada persona se esfuerza por maximizar su propia utilidad. Y al hacerlo trata de obtener todos los beneficios que pueda de los actos de los demás, y de desplazar todos los costos que pueda de sus propios actos a los otros. Pero también, trata de evitar aquellas interacciones con los demás que les proporcionen a éstos beneficios cuyos costos recaigan sobre él. En la interacción puramente natural, el deseo de evitar la interacción costosa no tiene fuerza independiente para restringir el deseo de beneficiarse haciendo caso omiso de los demás. Pero, en la interacción social, que exige la restricción voluntaria, el deseo de impedir la interacción costosa puede equilibrarse con el deseo de obtener beneficios y desplazar costos. La estipulación, al prohibir que unos saquen ventaja de los otros, representa la restricción más débil racionalmente aceptable para las personas que quieren evitar la interacción costosa con los demás, y la restricción más fuerte racionalmente aceptable para las personas que quieren tener la libertad de beneficiarse. Por lo tanto, la estipulación refleja la igual racionalidad de las personas que deben restringir su interacción natural a fin de establecer relaciones sociales mutuamente beneficiosas.

Como ya vimos, la interacción restringida por la estipulación le genera a cada persona un conjunto de derechos, que son los que esa persona aporta a la mesa de negociación de la sociedad como sus dotes iniciales. Le da un derecho a sí mismo, un derecho a los frutos de su trabajo y un derecho a esos bienes cuya posesión individual exclusiva es mutuamente beneficiosa, bienes que el individuo obtuvo inicialmente o mediante el intercambio. Si la posesión de esos bienes no fuera conveniente o no fuera posible, la estipulación le daría al individuo el derecho al uso de tales bienes en términos que internalizaran todos los costos. Sin esos derechos, las personas no estarían racionalmente dispuestas ni a aceptar la prohibición de utilizar la fuerza o el fraude necesarios para la competencia de mercado, ni a aceptar voluntariamente las estrategias y prácticas conjuntas necesarias para la cooperación.

5.2 Pero, ¿esta conciliación que hacemos de la razón y la

moral no parece demasiado buena para ser cierta? Recordemos el relato sobre amos y esclavos con que comenzamos este capítulo. Entonces presentamos la figura 9 (véase pág. 265) para ilustrar la situación y aquí agregamos la figura 10 para representar una variante significativa de ese relato. En ambas figuras I_n corresponde a la situación inicial de los amos y los esclavos, B_n representa el resultado del acuerdo como lo imaginaba el joven universitario, en el cual los ex esclavos continuaban sirviendo voluntariamente, B_c representa el resultado real en el cual los ex esclavos se niegan a aceptar la transferencia improductiva que implica continuar sirviendo, e I_c representa la situación hipotética, no cooperativa pero no coercitiva, determinada por la estipulación como la posición inicial negociadora. La diferencia entre las dos variantes del relato es evidente. Si los hechos ocurrieran como lo ilustra la figura 10, el resultado final para los ex amos, representado por B_c, les proporciona una expectativa de utilidad menor que la de la situación inicial I_n. Si la posición inicial negociadora queda restringida por la estipulación, los ex amos pierden con la cooperación.

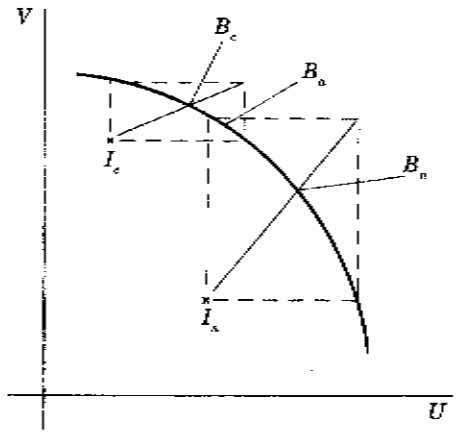


Figura 10

Sin duda, si los amos fueran racionales y capaces de prever el resultado, nunca aceptarían cooperar sobre bases que les resultarían desventajosas. No podemos suponer que la figura 10 ilustra una posibilidad real. Antes bien, debemos suponer o bien

que no se da la cooperación, con lo cual los amos y los esclavos permanecen en I_n, o bien la cooperación se da sobre bases que les proporcionen beneficios mutuos a ambas partes, con lo cual se llegaría a un punto B_a de la figura 10, en el cual los ex esclavos continúan sirviendo a sus ex amos lo suficiente para que los amos queden en mejor posición que en I_n.

Ninguna de las alternativas parece aceptable. La primera nos permite mantener firmemente nuestra afirmación de que es racional disponerse a acatar los acuerdos cooperativos sólo si éstos son (aproximadamente) justos y óptimos. Por ello, los esclavos no están dispuestos a aceptar ningún resultado que le proporcione a los amos cierta porción del superávit cooperativo. Debemos negar que pueda haber algún progreso racional en el resultado subóptimo no cooperativo. La segunda alternativa nos obliga a debilitar la estipulación y a fortalecer la disposición al acatamiento. Los derechos que tienen los amos en la posición inicial negociadora les permiten sacar alguna ventaja de los esclavos. Los esclavos están dispuestos a aceptar una estrategia conjunta que produzca un resultado óptimo que sea lo suficientemente justo para ser compatible con el objetivo de proporcionarles a ambas partes beneficios mutuos. Queda roto pues el vínculo entre la racionalidad y la imparcialidad. (Una tercera posibilidad que sería que los amos acepten las restricciones de la estipulación y cooperaran aun cuando perdieran en comparación con la no cooperación, debe quedar desechada porque es claramente incompatible con la concepción maximizadora de la utilidad de la racionalidad, aunque sea restringida.)

Para los maximizadores de la utilidad, el vínculo entre la cooperación y el beneficio mutuo debe tener prioridad respecto del vínculo entre la cooperación y la imparcialidad o la justicia. Pero eso no nos exige romper el segundo vínculo. La estipulación restringe la posición negociadora inicial hasta el punto, pero sólo hasta el punto, de que tal restricción sea compatible con el resultado cooperativo que le proporcione a cada individuo la expectativa de una utilidad mayor que la que le proporcionaría el resultado no cooperativo. Es racional acatar una estrategia cooperativa conjunta si y sólo si su resultado esperado es (aproximadamente) óptimo y lo suficientemente justo para ser compatible con el beneficio mutuo. Esto no significa que aban-

donemos ni la estipulación ni el acatamiento estrecho, sino que los subordinamos a la exigencia del beneficio mutuo.

Haremos la necesaria distinción entre el acatamiento, es decir, la disposición a aceptar los arreglos cooperativos justos y óptimos, y la *aquiescencia*, es decir, la disposición a aceptar los arreglos cooperativos algo menos que justos, a fin de asegurar el beneficio mutuo. Una persona que se dispone a la aquiescencia en una estrategia conjunta no piensa en la mayor utilidad que puede obtener en las ocasiones particulares en las que se adhiere a la estrategia. Restringe sus esfuerzos maximizadores de la utilidad. Pero no ignora la ventaja que concede a los demás al hacerlo y está dispuesta a rehusarse y a exigir una estrategia conjunta que le sea más favorable, si supone que una estrategia semejante ha de ser provechosa para los demás. Aquí suponemos, como sostiene Buchanan[22], que cada persona apela a la distribución natural o al resultado no cooperativo que está detrás de las realidades sociales existentes. A medida que ese equilibrio subyacente cambia, también cambian las condiciones de aquiescencia.

Por lo tanto, la cooperación basada en términos algo menos que justos es menos estable, pues no obtiene la aceptación sincera y plena de todos los participantes. La cooperación justa invita al consentimiento pleno gracias al cual nadie está dispuesto a rehusarse debido a los cambios en la distribución natural. Porque aunque nadie esté dispuesto a concederles ventaja a los demás a fin de lograr esa estabilidad, cada individuo desea, cuando no le resulta costoso, aceptar los arreglos como una forma de aumentar los beneficios de las instituciones y prácticas cooperativas.

Suponemos que el vínculo entre la razón y la moral se ha aflojado a causa del papel que desempeñan tres consideraciones completamente diferentes que afectan la interacción. En primer lugar están los factores ideológicos: creencias que afectan los términos de la cooperación de modos que a veces son injustos con aquellos que comparten las creencias, pero típicamente, con aquellos que no las comparten. Si se acepta normalmente que el lugar de una mujer es el hogar, una mujer que no estime satisfactorio el oficio de ama de casa, debería considerar imposible entrar en interacción con otras personas

de un modo que le proporcione una porción justa del superávit cooperativo. En segundo lugar, y con frecuencia relacionados con la ideología, están los recursos históricos: instituciones y prácticas que determinan efectivamente los derechos que las personas pueden ejercer en la interacción y si éstos satisfacen o no la estipulación. Si es habitual pagarles a las mujeres salarios más bajos que a los hombres, ninguna mujer en particular podría razonablemente esperar que se le reconozca que merece la misma paga por el mismo trabajo.

En última instancia, estos factores pueden depender de creencias irracionales; aunque la irracionalidad pueda estar firmemente incorporada a las prácticas e instituciones que los individuos no pueden efectivamente ignorar. Pero esa irracionalidad las hace menos perturbadoras para nuestra agumentación destinada a vincular la razón y la moral que el tercer factor, el factor tecnológico. La tecnología es el poder; aquellos que poseen una tecnología más avanzada con frecuencia están en posición de dictar los términos de la interacción a los demás. Sin sus armas de fuego, la pequeña cantidad de conquistadores españoles nunca habría podido derrotar a las civilizaciones indias americanas. Seguramente, los españoles preferían dominar antes que cooperar, pero está claro que la superioridad tecnológica que poseeían les aseguraba que ninguna cooperación debía proceder de una posición inicial que violara la estipulación.

Una tecnología superior le permite a su poseedor mantener racionalmente (y obliga a los demás a la aquiescencia de) acuerdos que, en clara violación a la estipulación, se basan en derechos diferenciados. Y no podemos suponer, como en el caso de la ideología y de sus efectos institucionalizados, que la base de esa tecnología sea irracional. Por el contrario, constituye la aplicación más plenamente racional de las creencias a la práctica. Podemos decir que aquellos que poseen una tecnología superior son más racionales que sus semejantes, porque son más capaces de relacionar e ingeniar medios para sus fines. Y esto revela por qué los factores tecnológicos debilitan el vínculo entre la racionalidad y la imparcialidad. Nuestro argumento en apoyo de la estipulación y en defensa del acatamiento estrecho se basa en el supuesto de que existe una igual racionalidad entre

las personas, que sin embargo las diferencias de la tecnología niegan.

No obstante, suponemos que la racionalidad desigual a la que se llega mediante las diferencias tecnológicas que existen entre las sociedades es accidental. Que no refleja diferencias subyacentes de capacidad de las personas que constituyen aquellas sociedades. Al liberarse de esa cosmovisión falsa y de las prácticas e instituciones a las que da lugar tal cosmovisión falsa, los seres humanos tienden a la igualdad tecnológica. Tienden pues hacia un estado de cosas en el cual la estipulación, el principio de concesión relativa minimáxima y la disposición al acatamiento estrecho, participan plenamente en la tarea de vincular las exigencias de la razón con las demandas de la imparcialidad o de la moral.

Al conciliar la razón con la moral, no pretendemos que nunca sea racional que una persona saque ventaja de otra, que nunca sea racional ignorar la estipulación, que nunca sea racional aceptar prácticas injustas. Semejante pretensión sería falsa. Afirmamos que la justicia, la disposición a no sacar ventaja de los semejantes, es la virtud que conviene a la cooperación, voluntariamente aceptada por personas igualmente racionales. La moral surge en y del acuerdo racional entre seres iguales.

Notas

1 La expresión "estipulación lockeana" procede de *Anarchy, Satate, and Utopia* de Robert Nozick, editado en Nueva York en 1974. El tema se trata en las págs. 175-182. John Locke establece su estipulación en *Two Treatises of Government,* Londres, 1690, en el segundo tratado, capítulo 5, párr.27-33.

2 Hobbes, *Leviathan*, cap. 14, pág. 64.

3 Véase J. M. Buchanan, *The Limits of Liberty: Between Anarchy and Leviathan*, Chicago, 1975.

4 *Ibid.*, págs 23-25.

5 *Ibid.*, pág. 24.

6 *Ibid.*, pág. 75.

7 *Ibid.*, pág. 79.

8 Buchanan afirma que la posición negociadora inicial "no puede clasificarse apropiadamente como una estructura de *derechos*, puesto que no se ha establecido ningún acuerdo formal" (*ibid.*, pág. 24). Pero consideramos que los derechos son aquello que proporciona la base antes que el objeto del acuerdo; véase 4.3 *infra*.

9 Véase J. F. Nash, "Two-person Cooperative Games", *Econometrica*, 21, 1953; Luce y Raiffa, págs. 140-143: J. C. Harsanyi, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge, 1977, págs. 167-169.

10 Con todo, Harsanyi distingue situaciones con o sin presiones amenazadoras. Véase *Rational Behavior*, pág. 168.

11 Véase Locke, segundo tratado, capítulo 5 y primer tratado, capítulo 9.

12 Locke, segundo tratado, capítulo 5, párr. 33.

13 Nozick, *op. cit.*, pág. 175.

14 Véase *ibid.*, págs. 57, 63-65.

15 Véase *ibid.*, pág. 185.

16 J. Rawls, *A theory of Justice,* Cambridge, Mass., 1971, pág. 74.

17 *Ibid.*, págs. 101-102.

18 *Ibid.*, págs. 101, 74.

19 Winch, *op. cit.*, pág. 99

20 Nozick, *op. cit.*, pág. 9.

21 Rawls, *op. cit.*, pág. 4.

22 Véase Buchanan, *op. cit.*, págs. 74-82, especialmente pág. 79.

VIII

El punto de Arquímedes

1.1 "Profundamente ligado a los principios de justicia hay un ideal de la persona que proporciona un punto arquimédico para juzgar la estructura básica de la sociedad."[1] La teoría moral ofrece un punto arquimédico para analizar la interacción humana. Pero, ¿qué es un punto arquimédico? El lector recordará que Arquímedes suponía que teniendo una palanca suficientemente larga y un punto de apoyo podía mover el mundo. Podemos entonces imaginar que un punto arquimédico es un punto de apoyo desde el cual un solo individuo podría ejercer la fuerza suficiente para mover o afectar cierto objeto. En la teoría moral, el punto arquimédico es aquella posisión que debe ocupar una persona, si sus decisiones tienen la fuerza moral necesaria para gobernar la esfera moral. Desde el punto arquimédico un individuo tiene la capacidad moral de dar forma a la sociedad. Caracterizamos dicho punto atendiendo a sus ocupantes, porque son las características intrínsecas de cada persona y no sus circunstancias ni sus artificios exteriores lo que le dan dicha fuerza moral.

¿Qué ideal de persona expresa el punto arquimédico? ¿Cómo conduce ese ideal a un juicio de la estructura de la sociedad o de los principios básicos que son el fundamento de la interacción social? y ¿qué importancia tiene ese juicio para la razón y la moral? Brevemente, suponemos que el ideal presenta un actor racional liberado, no de la individualidad sino del contenido de cualquier individualidad particular, un actor consciente de que es un individuo con facultades y preferencias, particulares en sí mismas y distintivas respecto de las de los demás, pero que no reconoce cuáles son esas facultades y preferencias. Un individuo semejante debe manifestar interés por las interacciones que mantiene con los demás y ese interés ha de llevarlo a elegir entre las estructuras sociales posibles. Pero su interés es necesariamente imparcial, porque se basa en los rasgos formales de la capacidad racional individual sin el

contenido tendencioso de un conjunto particular y determinante de características individuales.

Podemos suponer que esa persona ideal realiza sus elecciones en dos niveles. Desde el punto de vista abstracto, elige entre los principios de interacción; desde un punto de vista concreto elige entre las estructuras sociales que encarnan tales principios. Puesto que suponemos que el actor ideal es plenamente racional y que se sitúa en una perspectiva imparcial, creemos que su elección de principios expresa las normas de la justicia. Puesto que suponemos que es una persona plenamente informada (no por supuesto sobre sus propias facultades y preferencias, sino sobre las facultades, preferencias y circunstancias humanas), creemos que su elección de estructuras sociales debe ser la elección de una sociedad justa. Articular el razonamiento de esta persona equivale a revelar, como un sistema internamente racional, la moral de las personas que no se interesan por los intereses de los demás. Su elección expresa lo que sea que la moral les exige a las instituciones y prácticas de la sociedad.

Pero su elección expresa también las exigencias que le hace la igual racionalidad de todas las personas a la sociedad, concebida como la estructura de interacción en la que nació cada individuo. Aunque la sociedad sea un instrumento de sus miembros, no sólo es anterior a ellos sino que además influye en gran medida en la realización de sus facultades y en la naturaleza de sus preferencias, de modo tal que determina en parte el uso que cada uno de sus miembros quiera hacer de ella. Si nos preguntamos pues qué estructura social, qué instituciones y prácticas concitarían el acuerdo racional de todos los individuos como la mejor manera de determinar las condiciones sociales de la convivencia, encontraríamos la respuesta en la elección de un actor situado en el punto arquimédico. Porque, como lo mostraremos luego, esa elección refleja todas las facultades y todas la preferencias del único modo posible previo a la determinación de los individuos reales y particulares. Por consiguiente, podemos identificar la capacidad moral para dar forma a la sociedad con la capacidad de determinar el contrato social del que deben participar todas las personas racionales.

En capítulos anteriores analizamos la elección estratégica

en la perspectiva del actor individual. Luego extrajimos un principio de interacción cooperativa, el principio de la concesión relativa minimáxima, de las consideraciones maximizadoras generales y de una explicación de la estructura formal de las situaciones estratégicas. Mostramos que las personas racionales han de disponerse a acatar las restricciones a la maximización directa de la utilidad que establece ese principio, siempre que las dotes o derechos que las personas aporten a la interacción estén restringidas por otro principio, la estipulación lockeana. Por consiguiente hicimos derivar las restricciones internas a la maximización de la utilidad dentro del marco de la elección maximizadora misma. Identificamos los principios así establecidos con las exigencias de la justicia, demostrando que eran imparciales, siempre desde el punto de vista del individuo comprometido en la interacción estratégica.

Ahora partimos de la imparcialidad, caracterizando un punto de vista que expresa directamente sus exigencias. Atendiendo a esa perspectiva moral empleamos pues el artificio de la elección racional para lograr una derivación directa de los principios encarnados en las normas de la interacción justa. Esos principios expresan las exigencias de la igual racionalidad que deben cumplir las instituciones y prácticas sociales concebidas como algo anterior a los actores individuales quienes las utilizan como marco de sus interacciones. Si la moral y la racionalidad están en armonía, esos principios deben ser pues aquellos mencionados en nuestro primer análisis. La elección arquimédica debe seleccionar la maximización de la utilidad esperada individual, restringida por la estipulación y por la concesión relativa minimáxima. La perspectiva del actor ideal debe ser coherente con las perspectivas de los individuos racionales realmente comprometidos en la elección estratégica.

En la siguiente subsección ampliaremos nuestro enfoque del actor ideal y lo diferenciaremos de otro tipo posible de personas. Luego analizaremos los dos principales enfoques de la elección que se oponen a la idea del punto arquimédico, la teoría utilitarista de John Harsanyi y la teoría kantiana de John Rawls y mostraremos qué razones tenemos para rechazarlas.[2] Luego desarrollaremos nuestra propia explicación positiva y estableceremos la plena coherencia que existe entre la perspectiva ideal que proporciona el punto arquimédico y las perspectivas particulares de los actores racionales individuales. Con esto, la teoría de una moral basada en el acuerdo quedará plenamente justificada.

1.2 Apelar a la elección de una persona ideal, consciente de su individualidad pero no del contenido particular de esa individualidad, refleja cierta concepción de la empresa moral y del ideal de imparcialidad en los que se basa. La moral se ocupa de actores, de personas que se supone actúan y eligen y se supone, están implicadas en las consecuencias de sus hechos y sus elecciones. En su condición de actor, la persona ideal es un maximizador de la utilidad que busca obtener el mayor beneficio. Sus utilidades pueden ser independientes de las utilidades de los demás; el punto arquimédico es compatible con el supuesto de mutuo desinterés. Lo ideal queda determinado en gran medida por lo que está ausente antes que por lo que está presente; por supuesto el actor ideal es racional y generalmente está informado, pero busca su mayor utilidad sin cobrar conciencia de las circunstancias particulares en las cuales actúa, las facultades particulares que le permiten actuar y la preferencias particulares por cuya satisfacción actúa.

Pero la ignorancia del actor ideal se extiende únicamente a su incapacidad para indentificarse con una persona particular dentro de la sociedad. En cuanto a la naturaleza de la sociedad y de sus miembros el conocimiento que tiene el actor ideal es lo más completo posible. Conoce los límites y las variaciones de la capacidad y de los intereses humanos. Sabe cuáles son las estructuras sociales factibles y las funciones individuales que le asigna cada una. Sabe de qué modo afectan las estructuras sociales y las circunstancias particulares la manifestación de las capacidades y los intereses. Por lo tanto sabe cómo puede encajar las personas, definidas por sus capacidades y utilidades, y las sociedades, definidas por sus estructuras y funciones. Pero el actor ideal sólo posee ese conocimiento y la capacidad de utilizarlo en la elección como un ocupante del punto arquimédico. Como en el mito de Er, una vez que eligió debe beber las aguas del Leteo.[3] Pero a diferencia de aquellos que están ante Lachesis, elige no una vida, sino antes bien una sociedad, ignorando la vida que ha de encontrar en ella, y elige no basándose en

preferencias anteriores, porque éstas podrían darle una identidad particular en la elección, sino sólo considerando las formas en las cuales las preferencias, las facultades y las circunstancias pueden encajar dentro de las sociedades factibles.

¿Es posible semejante elección? Y si es posible, ¿puede tener alguna base racional? Debemos mostrar que la preocupación del actor ideal por maximizar las utilidades de la persona que él resulte ser proporciona una base determinada para elegir desde el punto arquimédico. Además, debemos mostrar que esa preocupación le permite a cada persona identificarse con el actor ideal. Porque aunque el hecho de ignorar la propia identidad excluye cualquier manifestación de parcialidad positiva en la elección, es insuficiente para garantizar una igual racionalidad. La imparcialidad de la elección se encuentra no en ausencia de la preocupación por aquellos que resulten afectados, sino en presencia de una misma preocupación, y esta última queda asegurada por el objetivo maximizador del actor ideal. Aunque ese actor ideal pueda no identificarse con nadie, todos han de identificarse con él. Por consiguiente, la imparcialidad del actor ideal queda manifestada en el carácter plenamente representativo de su elección.

A fin de afinar nuestra explicación podemos distinguir el actor ideal de otras personas ideales anteriormente consideradas relevantes para la moral, pero cuyos títulos rechazamos. Primero debemos observar que al desplazarnos hacia el punto arquimédico en ningún momento abandonamos la idea de Hume de que la razón es y debe ser la esclava de las pasiones. El ideal kantiano de una razón pura que es práctica a pesar de su completa indiferencia ante la pasión es completamente ajeno a nuestro enfoque.[4] Los imperativos de la razón continúan siendo dogmáticos. Y continúan siendo imperativos de la razón individual. No necesitamos analizar si para Kant es justo considerar en su teoría que existe un profundo conflicto entre la racionalidad y la individualidad, puesto que afirmamos que no hay afinidad entre nuestro enfoque y su visión de la razón práctica. Pero tenemos que señalar que el punto arquimédico anula los efectos deformantes de la individualidad, sin dejar por ello de sostener la idea, no sólo de la elección individual, sino también de la elección hecha por un individuo preocupado por

defender sus propios intereses. Rawls supone que sigue a Kant al tratar los principios de justicia "independientemente de las contingencias de la naturaleza".[5] En nuestra perspectiva, la individualidad humana no puede separarse de esas contingencias y los principios morales no pueden negarlas sino que deben reflejarlas.

Además del ideal kantiano expresado en el yo nouménico,[6] debemos presentar tres otras personas ideales: el observador ideal, el simpatizante ideal y el que podría llamarse el propietario ideal. El primero de ellos se caracteriza por su absoluto desapego de las estructuras sociales que evalúa y entre las cuales debe elegir.[7] En su carácter de observador puro, registra lo que se puede observar y lo registra correctamente. Registra los valores y las evaluaciones de aquellos a quienes observa. Pero ese registro no es una evaluación en sí misma. Su desapego le impide tener una base para la elección. Si suponemos que esa persona ha de elegir los principios de interacción que maximicen las utilidades alcanzadas por aquellos que se ajustan a tales principios, le estamos agregando a su elección una norma que no está presente en la observación misma. La observación es neutral entre la satisfacción y la frustración, entre la maximización y la minimización de la utilidad.

El simpatizante ideal es el personaje por completo opuesto al observador ideal: está totalmente comprometido en lugar de absolutamente desapegado. Pero los afectos de los demás sólo lo afectan hasta cierto punto. No tiene una preocupación personal por elegir entre principios de interacción o entre estructuras sociales que los encarnen. Sólo tiene una base para la elección en la medida en que se interesa por los intereses de los demás —lo cual por supuesto es central para la concepción de un simpatizante—. Pero esto es exigirle a la persona ideal lo que no necesariamente tiene que estar presente y no está supuesto en su sí mismo real. Los principios de justicia proporcionan una base moral para la interacción entre personas que no se interesan por sus intereses recíprocos; hacer derivar tales principios de la idea de una persona que necesariamente está interesada en sus semejantes sería fundamentalmente incoherente. Las exigencias morales que puedan hacérseles a las personas deben ser directamente inteligibles para esas personas, independien-

temente de que aquellas sean también exigencias racionales. Pero los principios establecidos desde el punto de vista de un simpatizante ideal carecerían de ese carácter inteligible; podrían parecer completamente ajenos a la interacción entre personas que no tienen el sentido de la simpatía. El punto arquimédico debe idealizar a la persona que se encuentra en las circunstancias de justicia.

Al rechazar al simpatizante ideal como la encarnación adecuada de la elección moral expresamos nuestro distanciamiento con Hume. Al rechazar la razón como base de la moral, Hume no descubre otra alternativa que considerar que esa base está en cierto sentimiento o afecto que le proporciona al individuo una identificación directa con los demás.[8] Coincidimos con Kant en cuanto a que la moral nos hace demandas que son y deben ser por completo independientes de cualquier sentimiento de solidaridad que podamos tener.[9] Por consiguiente debemos excluir cuidadosamente cualquier apelación a tales sentimientos cuando hablamos de la elección hecha desde el punto arquimédico.

Un propietario ideal, en el sentido en el que utilizamos el término, es alguien que tiene un derecho previo en la estructura social: un dios que puede determinar las condiciones de interacción que deben satisfacer sus criaturas. Pero la concepción misma de un propietario ideal, como la de un observador ideal no ofrece ninguna norma para la elección. El propietario ideal se interesa solamente por su derecho de posesión, pero ese derecho no determina ninguna forma de interacción entre sus criaturas. Sin duda tenemos que agregar algo de benevolencia a nuestra explicación del propietario ideal, pero eso no forma parte de nuestro análisis. Las demandas de una moral teísta, en su forma pura, deben resultarles a los seres que se consideran a sí mismos actores racionales e independientes, como dice Hobbes: "aguijones contra los que no se puede cocear".[10] Sólo el poder del propietario hace que sus demandas sean relevantes para los maximizadores de la utilidad y nosotros, a diferencia de Hobbes, rechazamos que la fuerza pueda ser la base del derecho.

2 ¿Cómo podemos caracterizar la elección hecha desde el punto arquimédico? Una primera sugerencia, adaptada de la teoría moral de John Harsanyi, es que esa elección implica una

extensión simple del procedimiento que se utiliza para hacer una elección natural en condiciones de riesgo e incertidumbre.11 Lo peculiar del punto arquimédico es sólo la forma de la incertidumbre: lo único que el individuo encuentra incierto al elegir es su propia identidad y no las circunstancias que lo rodean. Un individuo debe seleccionar principios básicos de interacción o estructuras sociales que los encarnen sin conocer ni la función que él ha de cumplir ni las características personales que han de definirlo en la estructura elegida. Aunque esto no complica la elección desde el punto de vista teórico, puede dificultarla en la práctica.

La elección del actor ideal debe ser imparcial. Veamos qué dice Harsanyi sobre cómo debe aplicarse esa imparcialidad. Harsanyi afirma que

> ...esa exigencia de impersonalidad e imparcialidad siempre podría satisfacerse si el individuo que elige tiene que hacerlo entre las...alternativas *sin saber* —o por lo menos voluntariamente desatendiendo— cuál ha de ser su propia posición social personal en el sistema resultante. Más específicamente, esas exigencias podrían satisfacerse si el individuo tuviera que elegir entre las ... alternativas suponiendo que tiene la *misma probabilidad* de ocupar cualquiera de las posiciones sociales existentes, desde la más encumbrada hasta la más baja. (Llamaré a este modelo el *modelo de equiprobabilidad* de los juicios de valor moral.)
>
> De acuerdo con la moderna teoría de la decisión, un individuo racional que se encuentra en esa situación de elección hipotética siempre habría de elegir la alternativa que le brinde la mayor *utilidad esperada*, la cual, según este modelo, equivaldría a elegir la alternativa que le brinde *el nivel de utilidad promedio* de los miembros individuales de la sociedad. Por consiguiente...hacer un juicio de valor moral implica tratar de maximizar el punto medio aritmético de todas las utilidades individuales.[12]

Harsanyi supone que, aunque la imparcialidad impide saber quién es uno, el individuo sabe quiénes son las personas de cada una de las estructuras sociales entre las cuales debe elegir. El individuo puede determinar para cada una de tales personas una utilidad esperada; el rendimiento que podría esperar dados los principios de interacción encarnados en la

sociedad. Al tener una absoluta incertidumbre sobre su propia identidad, uno debe suponer que si tuviera que elegir una sociedad particular, tendría una posibilidad igual de ser cualquiera de los miembros de esa sociedad. Por consiguiente, el individuo evalúa cada sociedad pesando la utilidad esperada de cada persona, mediante la probabilidad $p = 1/n$, en la que n representa la cantidad de personas que conforman la sociedad; luego el individuo suma las utilidades esperadas sopesadas para obtener una única utilidad esperada, que, por supuesto, es el promedio de las utilidades esperadas de los miembros de la sociedad. El individuo elige entonces aquellos principios de interacción, o aquella estructura social, que le proporcionen el promedio mayor de utilidad esperada. El individuo elige de acuerdo con el principio de utilitarismo promedio. Una vez más aparece la alternativa utilitarista, ahora como una forma de explicar la elección arquimédica.

Harsanyi supone que las utilidades de las diferentes personas son comparables, de modo tal que puede determinarse un promedio significativo. Nosotros deberíamos objetar esta suposición, pero en lugar de hacerlo sugeriremos una forma posible de hacer comparaciones de utilidad. Porque sostendremos que aun cuando pudiera determinarse un promedio significativo de utilidad, maximizar ese promedio no es, como afirma Harsanyi en el pasaje citado, maximizar la utilidad esperada de un individuo racional que no conoce sus preferencias ni el lugar que ocupa en la sociedad, no es por lo tanto, maximizar la utilidad esperada del actor ideal situado en el punto arquimédico. Por consiguiente, sostendremos que Harsanyi se equivoca al suponer que su procedimiento de elección es una simple extensión de la elección racional realizada en circunstancias de incertidumbre.

¿Cómo pueden compararse entonces las utilidades de las diferentes personas? En 2.3.2 presentamos una manera adecuada de asignar las utilidades numéricas que representan las preferencias de un individuo sobre los elementos de cualquier conjunto de resultados posibles: le atribuimos una utilidad 1 al resultado más preferido y una utilidad 0 al menos preferido e igualamos la utilidad de cualquier otro resultado Y con la utilidad esperada de una lotería L que obrara sobre los resultados más preferido y menos preferido, de modo tal que L fuera indiferente a Y. Sin embargo podemos suponer que le asignamos una utilidad 1, no al más preferido de los resultados posibles, sino a ese resultado hipotético (que sólo raramente constituye una posibilidad real) que pudiera ser completamente satisfactorio, es decir que satisficiera el máximo conjunto de deseos, objetivos e intereses del individuo, de modo tal que sería el más preferido entre los resultados lógicamente posibles. Y supongamos que correspondientemente le asignamos una utilidad 0 a ese resultado hipotético que pudiera ser completamente frustrante, de modo tal que sería el menos preferido entre los resultados lógicamente posibles. En ese caso deberíamos utilizar loterías sobre esos resultados para determinar las utilidades de todos los resultados realmente posibles. Medida en esa escala, la utilidad de cualquier resultado indicaría hasta qué punto tal utilidad satisface las preferencias de un individuo. Tendríamos pues una medida de la satisfacción de las preferencias generales.

Ahora las utilidades de las diferentes personas parecen completamente comparables. Se puede establecer la mayor utilidad, el mayor grado de satisfacción de la preferencia, si uno compara las utilidades de una sola persona o las utilidades de diferentes personas. Si determinamos las utilidades esperadas de cada persona en una estructura social factible siguiendo esta escala, el promedio de las utilidades esperadas de las personas que están dentro de cualquier estructura social, ha de representar pues el nivel promedio de la satisfacción de la preferencia en esa estructura. Podemos entonces identificar el utilitarismo promedio con la maximización del nivel promedio de la satisfacción de la preferencia. En la elección hecha desde el punto arquimédico, la maximización de la utilidad esperada en situaciones de incertidumbre se extiende simplemente a la maximización de la utilidad esperada promedio, en donde la utilidad esperada de cada individuo es una medida de hasta qué punto quedan satisfechas sus preferencias.

Si el argumento de Harsanyi fuera acertado, el actor ideal, al maximizar la utilidad promedio, maximizaría pues la utilidad esperada de la persona que resultara ser, puesto que la utilidad esperada de cada persona ignorante de la identidad que tiene en la sociedad, es simplemente el promedio de la utilidad

esperada. Cada persona podría entonces identificarse con el actor ideal; cada uno consideraría que su propia utilidad esperada queda maximizada en las desusadas circunstancias de la elección arquimédica. Del mismo modo, al representar a cada persona racional, el actor ideal determinaría los términos del contrato social mediante su elección.

Pero el argumento no es acertado. Consideremos la elección propuesta. Puede caracterizarse cada estructura social como una lotería en la cual todos los premios tienen las mismas probabilidades y cada premio resulta ser una persona con determinadas facultades y preferencias, que se encuentra en una particular situación social. Podemos representar tales premios por un par ordenado (a,x) en el que a es una determinada persona y x una situación particular. Le asignamos una utilidad —la utilidad esperada para una persona a en una situación x— al premio (a,x) y podemos representar esa utilidad como $U_a(x)$. Harsanyi pretende que para cualquier persona i, desconocedora de las características que ha de tener y de la situación social que ha de ocupar, la utilidad esperada de la estructura social es el promedio de las utilidades de los premios. Si una estructura social está formada por las personas a, b y c que se hallan respectivamente en las situaciones x, y, z, la utilidad esperada de esa estructura para cada persona i es $1/3$ $[U_a(x) + U_b(y) + U_c(z)]$.

Pero esto evidentemente es falso. Cuando una persona i considera el premio (a,x), ha de asignarle, no $U_a(x)$, que es la utilidad que obtiene la persona a estando en x, sino $U_i(a,x)$, es decir la utilidad que obtiene la persona i si la persona a estuviera en la situación x. La utilidad esperada de una estructura social simple presentada en el párrafo anterior es $1/3$ $[U_i(a,x) + U_i(b,y) + U_i(c,z)]$. Para que el argumento de Harsanyi se sustente, debemos suponer que todas las utilidades son comparables interpersonalmente, $U_a(x) = U_i(a,x)$, para todas las personas i y a, y todas las situaciones x. Pero no necesariamente es así. Lo sería si y sólo si el grado de satisfacción de las preferencias que obtiene i estando a en x es el mismo que obtiene estando ella misma en x.

Consideremos, para retomar un ejemplo hecho famoso por John Stuart Mill, a Sócrates insatisfecho y al necio satisfecho.[13]

La insatisfacción de Sócrates refleja el bajo nivel de su situación medido con la vara de sus "sabias" preferencias y la satisfacción del necio refleja el alto nivel de su situación medido con con la vara de sus "necias" preferencias. Pero, sabios o necios, no podemos incorporar a la medida interpersonal de la preferencia ninguna evaluación previa; si son las consideradas verdaderas preferencias de las personas interesadas deben evaluarse por lo que representan para cada persona. De ahí que debamos asignarle una mayor utilidad al necio en su situación que a Sócrates en la suya; U_{necio} *(satisfecho)* debe ser mayor que $U_{Sócrates}$ *(insatisfecho)*. Se sigue de ello, lector mío, ¿que usted debe preferir ser un necio satisfecho que un Sócrates insatisfecho o, en otras palabras, que U_{lector} *(necio, satisfecho)*, debe ser mayor que U_{lector} *(Sócrates, insatisfecho)*? Mill no deja ninguna duda: él preferiría ser un Sócrates insatisfecho que un necio satisfecho. Probablemente no todos coincidan, pero la preferencia de Mill es claramente racional.

Al elegir entre el Necio y Sócrates, Mill tiene en cuenta, no sólo las preferencias de ambos y hasta qué punto quedan satisfechas, sino también sus propias preferencias por tener ciertas características personales y hasta qué punto podrían quedar satisfechas. Harsanyi en cambio no tiene en cuenta estas últimas preferencias.[14] Pero sin duda son importantes. Al elegir como un utilitarista-promedio, debemos suponer que el actor ideal selecciona el "valiente nuevo mundo" de Aldous Huxley, en el cual las preferencias de cada persona encajan perfectamente con su función social (independientemente de los errores de adaptación particulares necesarios para generar una novela interesante), de modo tal que quede garantizada la máxima satisfacción de la preferencia. Si todo lo que importara fuera la satisfacción de la preferencia, deberíamos respaldar esa elección. Pero no lo hacemos. Tenemos que llegar a la conclusión de que el actor ideal, si maximiza la utilidad promedio, no necesita maximizar la utilidad esperada de cada individuo o quizá de ningún individuo ignorante de la identidad que tiene en la sociedad, puesto que el actor ideal no puede considerar las preferencias que tienen las personas atendiendo a sus posibles identidades.

Podría decirse que nuestro argumento muestra no el fraca-

so de la explicación utilitarista-promedio de la elección arquimédica, sino nuestra propia incapacidad de captar la plena imparcialidad de esa elección. Supusimos que el actor ideal no necesita maximizar la utilidad esperada de ninguna persona que ignore cuáles son su identidad y su función en la sociedad, porque el actor ideal no puede tener en cuenta las preferencias por las diferentes identidades. ¡Pero una persona ignorante de su identidad no tienen tales preferencias! O, antes bien, sus preferencias sólo pueden ser el promedio de las preferencias de las diferentes personas, y él puede ser cualquiera de ellas. Una persona que tiene preferencias particulares y determinadas por ciertas características personales ya no está en posición de elegir con plena imparcialidad. La plena imparcialidad, como queda expresada mediante la elección arquimédica, exige que la persona sea ignorante , no sólo de las características que tendrá y las funciones que cumplirá en cada una de las estructuras sociales factibles entre las cuales debe elegir, sino también de las características de las personas que están incluidas en las estructuras sociales. El individuo debe suponer, pues, que existe la misma posibilidad de tener las preferencias de cualquier persona. Si una estructura social está formada por las personas *a, b, c* que se encuentran en las situaciones *x, y, z*, el individuo considera la estructura como una lotería en la cual si él es la persona *a* recibe el premio (a,x), si es la persona *b*, recibe el premio (b,y) y si es la persona *c*, recibe el premio (c,z). Si existe la misma posibilidad de ser cualquiera de esas personas, el individuo tiene las mismas posibilidades de recibir las utilidades $U_a(x)$, $U_b(y)$ y $U_c(z)$, de modo tal que su utilidad esperada es, después de todo, $1/3\ [U_a(x) + U_b(y) + U_c(z)]$. El actor ideal, al maximizar la utilidad promedio esperada, maximiza la utilidad esperada de cada persona que elige sin conocer ni la identidad ni la función que habrán de asignársele en la estructura social elegida, ni tampoco su propia identidad actual.

Pero hagamos una reflexión sobre la elección que le estamos adjudicando al actor ideal. Al no saber quién es, el actor ideal no puede expresar un único conjunto de preferencias que pueda ser representado por una única función de utilidad. No tiene un punto de vista único y unificado desde el cual puede establecer un orden de preferencias sobre las estructuras sociales entre las que elige. Establece un orden comparando la satisfacción de las preferencias correspondientes a los diferentes conjuntos, los cuales reflejan puntos de vista diferentes y posiblemente incompatibles. Pero ese orden no es el orden de preferencias de ninguna persona; no es un orden de preferencias en absoluto. Su medida, la utilidad promedio, es la medida de hasta qué punto se satisface la preferencia, pero puesto que no es la medida de un orden de preferencia, no es una utilidad.

El actor ideal, al maximizar la utilidad esperada promedio, no está maximizando la utilidad esperada, ni la propia ni la de ningún otro. Por consiguiente, la elección hecha desde el punto arquimédico no es una simple extensión de la elección racional hecha en situación de incertidumbre. Al colocar la incertidumbre en la base de la elección, al negar la unidad que proporciona un conjunto único de preferencias, el carácter de la elección cambia radicalmente. Pero aunque esto muestra que Harsanyi confunde la relación entre el principio utilitarista-promedio y el principio de maximización de la utilidad esperada, no muestra que Harsanyi haya caracterizado incorrectamente la elección hecha desde el punto arquimédico. Si aceptamos que esa elección es distintiva, ¿por qué no deberíamos suponer que el principio utilitarista-promedio captó ese carácter distintivo? Si para una persona ignorante de su propia identidad no es posible una simple extensión de la maximización de la utilidad esperada, puesto que una persona semejante carece del conocimiento de su propia función de utilidad, ¿por qué la maximización de la utilidad promedio no habría de recomendarse a sí misma como el sustituto apropiado, hecho a la medida de las circunstancias arquimédicas?

Al distinguir la utilidad esperada promedio de la expectativa de utilidad de cualquier individuo, quebramos el vínculo que garantizaba que cada persona se identificara con el actor ideal situado en el punto arquimédico. Al reconocer que la maximización de la utilidad esperada promedio exige elegir una sociedad de Necios satisfechos antes que una de Sócrates insatisfechos, podemos sospechar que es imposible encontrar otro vínculo que garantice esa identificación. En realidad, sostenemos que el enfoque utilitarista-promedio de la elección arquimédica no llega a comprender al individuo humano. Aun-

que supongamos que cada persona intenta maximizar su utilidad, no podemos decir que esa persona se considere a sí misma como un mero receptáculo de la satisfacción de la preferencia. Suponer que un individuo ha de elegir entre las estructuras sociales sólo por la maximización del nivel promedio de satisfacción de la preferencia es suponer, no que tal individuo ha de elegir imparcialmente, sino que ha de exhibir un completo desinterés por conocer su propia identidad, sus facultades y sus preferencias y que sólo ha de preocuparse por aquellas facultades que le permitan satisfacer sus preferencias.

Cada persona, consciente de su identidad, debe poder identificar la elección hecha desde el punto arquimédico como la elección que hubiera hecho en condiciones de acción ideales. Porque una elección es racional sólo para aquellos capaces de identificarse con ella y es imparcial sólo si todos pueden identificarse con ella. Pero la elección que exige el utilitarismo promedio no suministra una base para tal identificación. John Rawls declaró: "El utilitarismo no toma en serio la distinción entre las personas".[15] Estamos de acuerdo con él pero además agregamos que el utilitarismo no toma en serio la individualidad de las personas. Nuestra objeción no se refiere tanto a que una persona pueda sacrificar su utilidad por aumentar la de los demás, sino a que una persona sea tratada como un medio para lograr la satisfacción de la preferencia general. El utilitarismo viola la integridad del individuo como ser humano con sus propias facultades y preferencias características y por lo tanto, con una utilidad distintiva que no es intercambiable con las utilidades de los demás que él trata de maximizar.

Pero la idea básica de este análisis no es atacar directamente el utilitarismo. Aceptamos que si la elección hecha desde el punto arquimédico tuviera que representarse en términos directamente maximizadores, tal representación debería ser utilitarista. Si consideramos la elección arquimédica como una ampliación de la elección racional en circunstancias menos inusuales, debemos determinar un único orden de resultados posibles, en términos (de los cuales se pueda definir una medida) que puedan ser maximizados. Y el único orden que puede satisfacer las necesarias condiciones de imparcialidad e igual racionalidad es un orden que responda positivamente a los

órdenes de preferencia y que evalúe del mismo modo el orden de cada individuo. La demostración de Harsanyi, a la que nos referimos en 5.2.2, de que el bienestar social debe ser un promedio de las utilidades individuales, muestra que una elección maximizadora imparcial sólo puede ser utilitarista.

Por consiguiente, nuestra conclusión primaria debe ser que la elección hecha desde el punto arquimédico no es una extensión de la maximización de la utilidad esperada. Porque al no proceder de un conjunto de preferencias bien definido y unificado, adquiere un carácter completamente distintivo. No sólo no implica la maximización de la utilidad esperada de ningún individuo, sino que tampoco implica la maximización de ninguna medida definida de las utilidades individuales. ¿Qué implica entonces? ¿Cómo podemos representar la elección hecha desde el punto arquimédico, de modo tal que cada persona, consciente de sus propias facultades y preferencias, sea capaz de identificarla con la elección que haría si no fuese consciente de ellas? Ese es el problema que aun debemos resolver.

3 El actor ideal se preocupa por su identidad, sea ésta cual fuere. La elección de principios para la interacción social que haga debe reflejar tal preocupación. Debe vincularlo con cada una de las demás personas; cada individuo debe estimar el ideal como un ideal de sí mismo con todas las características particulares, y por lo tanto tendenciosas, excluidas de su deliberación. En la teoría de la justicia de John Rawls podemos encontrar un intento consciente de establecer ese vínculo. La idea de caracterizar a la persona ideal como a alguien situado detrás de un "velo de ignorancia" con cuyas elecciones todos pueden identificarse.[16]

Será útil para nuestro propósito analizar en su forma más general los conceptos de Rawls. Aunque Rawls se concentra en un caso especial, aplicable a las circunstancias de las sociedades desarrolladas actuales, en su forma más general sus conceptos deben defenderse pues vinculan, en el nivel más profundo, el punto arquimédico con los individuos reales. Su concepción general de la justicia es que "todos los valores sociales se distribuyan equitativamente, salvo que una distribución desigual de algunos de ellos, o de todos, sea ventajoso para cada uno de los miembros de la sociedad".[17] Esta idea está expresada en un principio fundamental que gobierna la interacción, "el prin-

cipio de diferencia lexicográfica", en el cual se basan los principios de justicia más específicos que constituyen el foco del análisis de Rawls. Este principio establece que "en una estructura básica con *n* representantes relevantes, primero se maximice el bienestar del representante peor dotado; segundo, en el caso de que otro representante tenga el mismo bienestar que el peor dotado, se maximice el bienestar del segundo peor dotado y así sucesivamente".[18] Antes de continuar con la idea de las personas representativas, observemos que Rawls la utiliza con el fin de excluir un preocupación directa por los costos y los beneficios que implican una relación accidental e impredecible con las condiciones de la interacción social, las cuales no formarían parte del enfoque referente a personas consideras representantes de las posiciones sociales.

El principio de diferencia lexicográfica exige que las estructuras sociales estén diseñadas con el fin de maximizar el nivel mínimo de utilidad; es decir que prescribe la utilidad maximínima. Está claro que este principio exige que haya utilidades interpersonalmente comparables; y podemos suponer que la comparación se basa en el grado de satisfacción de la preferencia. Pero el propio Rawls deja de lado el problema de las comparaciones; y en realidad los principios más específicos que analiza se refieren no a la utilidad sino a los "bienes primarios": los bienes que uno querría si quisiera algo.[19] Obsérvese que, en constraste con la exigencia de utilidad maximínima, el principio de concesión relativa minimáxima exige que las estructuras sociales estén diseñadas con el fin de maximizar el nivel mínimo de beneficio relativo. Sostuvimos antes que, desde el punto de vista de las personas comprometidas en una negociación, es racional ponerse de acuerdo para maximizar la mínima proporción de la ganancia de utilidad que cada uno pueda obtener de la cooperación. Rawls afirma que, desde el punto de vista arquimédico, es justo maximizar la cantidad mínima de utilidad, sin distinguir la porción de ella que constituye el excedente cooperativo.

¿Por qué habría de elegir el actor ideal el principio de diferencia lexicográfica? Aunque los argumentos de Rawls son complejos y están presentados de un modo discursivo e informal, podemos identificar por lo menos dos temas principales.[20] El

primero es que, al elegir ignorando la propia identidad, el individuo debe tener especial cuidado de evitar lo peor que pueda acontecer. El segundo es que los términos de la cooperación, en la medida en que beneficien a algunos más que a otros, deben hacerlo de un modo que concite el apoyo voluntario de aquellos menos favorecidos. Podríamos decir que aunque no lo afirme explícitamente, en el pensamiento de Rawls está presente la idea de que si el menos beneficiado puede identificarse con la elección arquimédica, todos pueden hacerlo y que por lo tanto, tal elección satisface las exigencias de igual racionalidad y de imparcialidad. Es evidente que esas dos consideraciones medulares —evitar lo peor y convocar la cooperación voluntaria de los menos favorecidos— defienden un principio maximínimo como el de la diferencia lexicográfica.

Rawls desarrolla largamente su argumentación cotejando sus principios de justicia con la concepción utilitarista expresada en la maximización del nivel promedio de utilidad.[21] Supongamos que uno se encuentre entre los menos favorecidos. Si los principios de la sociedad fueron elegidos sobre bases utilitaristas, los hemos de considerar de manera muy diferente que si hubieran sido elegidos para maximizar la utilidad mínima. En el primer caso, la persona menos favorecida atribuiría su propia situación desventajosa a una preocupación por la satisfacción de la preferencia que ignora el carácter distintivo y la individualidad de las personas. Se consideraría a sí mismo sacrificado por el conjunto de la sociedad. Pero en el segundo caso, la persona menos favorecida habría de reconocer que, aun cuando en alguna otra estructura social él podría hallarse mejor situado, ninguna alternativa podría impedir que alguien estuviera por lo menos tan desfavorecido como lo está él en la situación presente. Sin bien esto no es ningún consuelo para nadie, puesto que el individuo no se interesa por los intereses de los demás, no es racionalmente posible esperar que los principios de la propia sociedad le presten a ese individuo mayor consideración de la que le prestan a sus semejantes. El principio de diferencia lexicográfica garantiza que no se sacrifiquen los intereses de nadie a fin de mejorar la situación de algún otro. De ahí que si una persona se encuentra entre las menos beneficiadas, es capaz de identificarse con una elección basada en la diferencia

lexicográfica, aunque no con una elección basada en la maximización de la utilidad promedio.

Como una defensa del principio de diferencia lexicográfica contra la alternativa utilitarista, la apelación de Rawls a la posición del menos favorecido parece muy poderosa, pero no basta para hacer una demostración general del principio. Porque no muestra que el principio pueda concitar el apoyo de todos o que no haya una alternativa que también pueda atraer el apoyo de los menos favorecidos. Una defensa positiva de la posición de Rawls exige la apelación a una condición muy fuerte de beneficio mutuo: cada persona puede esperar que se satisfagan sus preferencias en la medida en que, pero sólo en la medida en que, esa satisfacción no obstaculice una expectativa igual de satisfacción de ninguna otra persona. El derecho que tiene cada persona al beneficio depende pues enteramente del efecto que pueda tener sobre lo que reciben los demás el hecho de que ella reciba tal beneficio. En particular, ninguna persona tiene derecho a ningún beneficio, salvo como un miembro de la sociedad y ninguna persona tiene derecho a ningún beneficio como un miembro de la sociedad en virtud de su contribución a la producción de bienes, ya sea en la interacción de mercado, ya sea en la interacción cooperativa.

Este último punto parece perder claridad cuando Rawls dice: "Aquellos que fueron favorecidos por la naturaleza, sean quienes fueren, sólo pueden obtener ganancias de su buena fortuna en términos que mejoren la situación de aquellos menos favorecidos. Los que fueron naturalmente beneficiados no deben ganar meramente a causa de haber recibido una mejor dote, sino sólo lo necesario para cubrir los costos de la formación y la educación y para utilizar sus dotes en formas que también ayuden a los menos afortunados".[22] Esto parece sugerir que aquéllos con la capacidad de ser más productivos desde el punto de vista social que sus semejantes pueden esperar obtener mayores beneficios, porque las recompensas serían incentivos que los llevarían a utilizar plenamente sus facultades y por lo tanto, al maximizar la productividad social, a beneficiar al conjunto de la sociedad y en particular a los menos favorecidos. Pero si se le presta la debida atención a lo que dice Rawls, se advierte que los más productivos no tienen derecho a una

porción mayor de los bienes en virtud de su mayor contribución. Si concederles una porción mayor aumentara su contribución y si esto, a su vez, elevara el mínimo social, ellos deberían recibir una porción mayor. Pero sus recompensas son puramente instrumentales, son sólo medios para alcanzar el objetivo de maximizar la utilidad mínima y no constituyen un reconocimiento del derecho de los más favorecidos, basado en su mayor contribución.

Rawls rechaza el concepto mismo de la contribución individual a la productividad social. Y afirma que "desde el punto de vista de la sociedad como conjunto *vis-à-vis* cualquier miembro, no existe ningún conjunto de objetivos acordados mediante el cual puedan evaluarse las contribuciones sociales potenciales de un individuo... La idea de una contribución individual a la sociedad considerada ésta como una asociación, se debilita".[23] Pero este argumento parece mal fundamentado. Porque, como el propio Rawls sostiene, la sociedad "es una empresa cooperativa destinada a la ventaja mutua",[24] y la idea de la ventaja mutua proporciona una norma para evaluar las contribuciones sociales de los individuos. Debemos por lo tanto considerar que Rawls sólo quiere decir que la contribución que hace un individuo no le da derecho a ninguna retribución de la sociedad. Y esto está implícito en el espíritu de Rawls cuando dice "El principio de la diferencia en efecto representa un acuerdo respecto de la distribución de los talentos naturales como un bien común y respecto de compartir los beneficios de esa distribución sea lo que fuere lo que ésta resulte ser".[25] Si los talentos constituyen un haber común, ni el talento de algún individuo ni el esfuerzo basado en ese talento, permiten abrigar ninguna pretensión particular a los productos de la interacción social.

Rawls también rechaza cualquier intento de dividir la utilidad de un individuo en componentes sociales y no sociales, o cualquier otra sugerencia de que un individuo tiene derecho a aquello que podría esperar independientemente de la sociedad más una porción de los beneficios adicionales que produzca la interacción social. Porque afirma que "nuestra naturaleza independientemente de la sociedad es sólo una base potencial con una amplia gama de posibilidades... independientemente del lugar que ocupemos y de la historia que tengamos en una

sociedad, ni siquiera podemos conocer nuestras aptitudes potenciales y aun nos falta formar nuestros intereses y nuestro carácter".[26] Las características particulares que individualizan a una persona y determinan sus preferencias son productos sociales; no proporcionan una base para satisfacer la preferencia independientemente de la sociedad, satisfacción que se suplementa mediante la interacción social. Al crear las preferencias, las características particulares del individuo limitan el *locus* de la satisfacción de la preferencia a la sociedad.

En la perspectiva de Rawls no parece suficiente que los principios de interacción sean imparciales, al limitarse a aumentar los intereses de cada persona sin acrecentar más los de unos que los de otros. Antes bien parece necesario que esos principios hagan abstracción de lo que Rawls considera un elemento moralmente arbitrario de la individualidad, de modo tal que tales principios consideren a las personas independientemente de aquellas características que son meramente particularizantes y por lo tanto fuentes potenciales de deformación, porque todas ellas son arbitrarias desde el punto de vista moral. Nadie tiene ningún derecho basado en el hecho de ser la persona particular que es, y nadie tiene semejante derecho porque las características particulares de una persona son en sí mismas los efectos de la interacción social. Cada persona, considerada independientemente de su lugar y su tiempo particulares es una mera potencialidad, y los principios de interacción justa sólo pueden tener en cuenta la individualidad de alguien como potencialidad y en la medida en que ésta no lo diferencie de las demás personas.

Ahora estamos en posición de elaborar el argumento para defender el principio de diferencia lexicográfica. Un individuo, inconsciente de su identidad, debe elegir los principios que han de encarnarse en la estructura básica de la sociedad. Ese individuo reconoce que sus características personales, sean las que fueren, están formadas dentro de esas estructuras sociales y que por lo tanto no le permiten pretender ni beneficios extrasociales ni ninguna otra porción del producto social. Su contribución a la sociedad, en la medida en que dependa de esas características formadas socialmente, tampoco le permiten pretender los beneficios obtenidos por la sociedad. De ahí que,

sea quien fuere este individuo, no puede tener ninguna pretensión particular a obtener los beneficios de la interacción social. Por consiguiente, puede esperar que los demás estén dispuestos a participar con él en la interacción sólo sobre bases que no le proporcionen a él ningún beneficio que prive a los demás de un beneficio igual. Del mismo modo, los demás pueden esperar que ese individuo esté dispuesto a participar sobre bases que no les proporcionen a los otros un beneficio que lo prive a él de un beneficio igual. Pero, puesto que sea quien fuere nuestro individuo, ha de buscar su máximo beneficio, esto le dicta que elija que las estructuras básicas de la sociedad encarnen el principio de utilidad maximínima a fin de regular la interacción social. Sea quien fuere, el individuo puede identificarse con esa elección, puesto que ésta representa la mayor demanda que él puede hacerles a los demás, en una situación en la cual, al ignorar sus características particulares, no puede hacer que sus demandas se ajusten a las habilidades y talentos deformantes que lo individualizan.

Esta argumentación no surge tan claramente en la obra principal de Rawls, *A Theory of Justice*. Allí Rawls escribe: "No hay más razón para permitir que la distribución de los ingresos y la riqueza se basen en la distribución de los haberes naturales que para permitir que se basen en la fortuna histórica y social... Porque una vez que nos dejamos perturbar por la influencia o bien de las contingencias sociales o bien de la suerte natural al determinar las porciones distributivas, estamos atrapados y siempre nuestra reflexión recibirá la influencia de la otra. Desde un punto de vista moral tanto las contingencias morales como la suerte natural son igualmente arbitrarias".[27] Este argumento, que admite la diferencia entre las características naturales y las contingencias sociales, mientras niega que esa diferencia tenga relevancia moral, no encaja bien con la idea de la elección arquimédica. Porque aunque el actor ideal sea inconsciente de su identidad, es consciente de que tiene una identidad. Por lo tanto parece racional que esa persona eligiera teniendo presente esto último y que considere que, dada esa identidad, puede pretender los frutos de la interacción social. Esa persona elegiría pues un principio destinado a regular la interacción, tal que las características naturales particulares de cada persona (en la

medida en que afecten lo que cada persona pueda hacer) sean tenidas en cuenta al determinar la distribución de los beneficios. Y esas características particulares constituyen los haberes naturales de la persona. Por consiguiente, deberíamos esperar, contrariamente a lo que sostiene Rawls, que en la elección arquimédica lo que debería tenerse en cuenta como parte de la identidad desconocida pero real del individuo, son los haberes naturales y no las contingencias sociales. Cada persona aporta a la sociedad sus haberes naturales; y éstos entran en la formación de las preferencias y determinan cuál es la contribución social que hace ese individuo. Y no parecen ser moralmente arbitrarios en ningun sentido que pudiera excluirlos como datos para ser tenidos en cuenta en la elección imparcial. Lo que sí excluye el punto arquimédico es que se ajusten los principios a fin de favorecer las propias dotes particulares. Pero, ¿ por qué habría de excluir que se ajusten los principios a fin de relacionar la distribución del producto social a la forma en que cada persona invierte sus dotes en la producción?

La respuesta, nunca dada en *A Theory of Justice*, pero que sí aparece en "The Basic Structure as Subject", es que los haberes naturales son en definitiva contingencias sociales.[28] Si al determinar las porciones distributivas nos sentimos perturbados por la influencia de los primeros, también debemos recibir la influencia de la suerte natural, porque lo que parece ser la fortuna o suerte natural es en gran medida una contingencia social. Puesto que los haberes naturales son productos sociales, dejar que sean éstos los que determinan la distribución social equivaldría a basar una contingencia social en otra. Independientemente de la sociedad, la naturaleza humana es meramente potencial; el hecho de que se traduzca en determinadas características individuales es el resultado de circunstancias históricas y sociales. Aunque esas características determinadas constituyan identidades individuales, tales identidades no suministran una base extrasocial para pretender los beneficios de la sociedad. El carácter ideal del actor situado en el punto arquimédico se debe por lo tanto a su incapacidad (a causa de que ignora su propia identidad) no sólo para adaptar los principios a su propia naturaleza y sus circunstancias, sino también y lo que es aun más importante, para suponer que su naturaleza

constituye un factor relevante al considerar cómo deben distribuirse los frutos de la interacción.[29]

Como ya hemos dicho, la justicia es la disposición a no sacar ventaja de los demás, ya sea como jugadores líberos, ya sea como parásitos. Aparentemente, el principio de diferencia lexicográfica autoriza a aquéllos con menos talento natural a sacar ventaja de aquéllos naturalmente mejor dotados, pues exige que estos últimos utilicen sus facultades, no primariamente para aumentar sus propio bienestar, sino para maximizar el nivel mínimo de bienestar. Pero, para poder criticar a Rawls en ese aspecto debemos suponer que cada individuo puede definirse independientemente de la interacción social, de modo tal que aporte sus talentos y aptitudes, sus actitudes y sus preferencias, a la sociedad como parte de una dote natural. Debemos suponer que las características particulares de cada persona, al permitirle hacer cierta contribución al producto social, también lo autorizan a pretender cierta porción de ese producto.

Esto es precisamente lo que rechaza Rawls y ese rechazo lo lleva a adoptar un enfoque muy diferente de lo que exige la justicia. La persona que saca ventaja de sus semejantes no es el individuo menos dotado, el cual se beneficia del principio maximínimo, sino el mejor dotado que utiliza su talento únicamente para aumentar su propio beneficio. Pues este último desvía para su uso exclusivo una proporción indebida del haber total de la sociedad. Les roba a sus semejantes lo que por derecho les corresponde.

Podríamos oponernos a la afirmación de Rawls de que independientemente de la sociedad los seres humanos no tienen una naturaleza determinada. En realidad, podemos suponer que Rawls exagera dramáticamente la influencia que ejercen las particulares circunstancias históricas y sociales en la formación de los seres humanos. Pero podemos soslayar las difíciles cuestiones fácticas implicadas en esto, cuestiones en las que no soy un experto. Aceptemos sencillamente que los seres humanos son productos sociales. ¿Debemos por consiguiente aceptar que la sociedad o cualquier grupo de personas vinculadas de algún modo, adquiere algún derecho a tales productos? El hecho de que las facultades de cada individuo se realicen dentro de un nexo social no hace que tales facultades se conviertan en

haberes colectivos antes que individuales, ni suministra ningún fundamento para considerarlas parte del capital común del grupo dentro del cual se ha formado el individuo. Al tomar para sí tales haberes, el individuo no se apodera de lo que por derecho les corresponde a los demás, puesto que él mismo no les pertenece a esos otros, aunque esto tampoco implica que el individuo pueda utilizarlos en su provecho por derecho propio.

Aceptamos, por lo menos a los fines de este argumento, que la génesis de la identidad individual es social. Pero no aceptamos que esa génesis afecte la condición de la identidad. Para ser uno mismo el individuo no necesita hacerse a sí mismo. Y el interés de cada persona por sí misma, su naturaleza de maximizador de la utilidad individual, no tiene necesariamente que verse afectado por el hecho de saber que son las fuerzas sociales las que dieron forma tanto a ese interés como a las preferencias determinadas que lo constituyen. Al participar con los demás en la interacción, cada persona considera que su propia naturaleza es sencillamente un don y no tienen ninguna razón para pensar de otro modo.

Según las palabras de Rawls que citamos antes, la sociedad es "una empresa cooperativa destinada a la ventaja mutua", pero no es simplemente un instrumento o una herramienta que podamos utilizar a fin de aumentar la eficacia de la interacción y que permita que cada persona mejore el modo de alcanzar sus propios fines. Considerar la sociedad como algo meramente instrumental equivale a pasar por alto tanto los efectos que produce la interacción social en la formación de los individuos y de sus fines, como el hecho de que las personas se interesan por algunos intereses de algunos de sus semejantes. Equivale a pasar por alto tanto la socialización como la sociabilidad. Pero tener en cuenta estos aspectos, aunque sea algo esencial si pretendemos evitar un instrumentalismo social parcial, no nos lleva a negar la función que cumple la interacción social en cuanto a ayudar (u obstaculizar) a cada individuo en la búsqueda de los fines que implican ausencia de interés por aquellos con quienes participa en la interacción, y no necesariamente afecta el modo en el que aceptemos regular tal interacción.

Cuando el actor ideal elige entre los principios factibles de interacción, lo hace, no lo olvidemos, en su carácter de actor. Su

elección es la elección de una persona con facultades, actitudes y preferencias particulares, aun cuando al elegir el individuo ignore cuáles son esas características. Pero si esa persona no tiene preferencias que le permitan ordenar los resultados posibles, no tiene capacidad para elegir entre las posibles acciones y no comprende la relación que existe entre las acciones como objetos de la elección y los resultados como objetos de la preferencia, no es un actor. Con lo cual no estaría implicado en la elección de los principios de interacción.

El actor ideal no necesita aceptar sus características particulares para saber que su racionalidad se expresa en el hecho de buscar el vínculo más eficiente entre elección y preferencia. Suponer que los factores que determinan la elección y la preferencia son arbitrarios es divorciar la moral de las condiciones de la acción racional. Que los principios de interacción social, para ser moralmente aceptables, deban ser imparciales entre esos factores, y por lo tanto deban elegirse de manera tal que ningún individuo pueda acomodarlos o deformarlos a fin de favorecer sus características personales particulares, no implica que tales principios den por descontadas o ignoren las bases de la elección y de la preferencia. No se alcanza la imparcialidad considerando que las personas carecen de facultades e intereses particulares; en realidad, semejante enfoque viola la imparcialidad entre los individuos reales. Al suponer que la distribución justa de los beneficios y los costos de la interacción social no debe relacionarse con las características de los individuos particulares que conforman la sociedad, Rawls viola la integridad de los seres humanos como son y como ellos mismos se conciben. Al tratar a las personas como seres puros liberados de la arbitrariedad de sus características individuales, Rawls llega a tratarlas como simples instrumentos sociales. Al negarle a cada persona el derecho a sus haberes naturales, Rawls termina por colectivizar tales dotes. En su argumentación la moral queda divorciada del punto de vista del actor individual.

Coincidimos con Rawls en que el utilitarismo no toma en serio la distinción entre las personas. Pero además creemos que el error más grave de esa corriente es no tomar en serio la individualidad de las personas. La teoría de Rawls cae en el mismo error porque la individualidad de las personas, conside-

radas éstas como actores, se basa en sus facultades y preferencias. La alternativa a maximizar la producción de la utilidad es no igualar la distribución, sino, antes bien, maximizar las porciones que le corresponden a cada individuo, porciones que son proporcionales a la contribución hecha por cada individuo que las recibe. Sólo así estimamos seriamente la individualidad de las personas, por socializadas y sociables que puedan ser, personas que sin embargo se encuentran en interacción con otras por las cuales no necesariamente deben sentir interés, personas que tratan de regular esa interacción de un modo con el cual todos puedan identificarse.

4.1 Nuestro análisis del principio utilitarista de maximización de la utilidad promedio y del principio generalizado de maximización de la utilidad mínima de Rawls nos prepara para elaborar un análisis constructivo de la elección que hace el actor ideal de los principios de interacción y las estructuras sociales. En el corazón mismo de ese análisis está la relación que existe entre la persona ideal, consciente de que es un actor, pero inconsciente de su identidad particular, y las diferentes personas reales, una de las cuales en realidad es ella, y cuyas interacciones quedarán reguladas por su elección. Hemos afirmado, de acuerdo con Rawls, que cada individuo debe poder identificarse con el actor ideal, de modo tal que cada uno pueda reconocer en la elección hecha desde el punto arquimédico, la elección que podría haber hecho él si hubiese estado sujeto a las restricciones impuestas al conocimiento que definen ese punto. Por consiguiente, el actor ideal debe elegir, no como si tuviera las mismas posibilidades de ser una de las personas afectadas por su elección, sino como si fuera realmente una de esas personas. Podemos concebir el punto arquimédico como un punto de convergencia; partiendo de cada individuo, que elige desde su propia perspectiva principios de interacción social —principios que por supuesto han de reflejar el interés de cada elector por maximizar su propia utilidad— alteramos las perspectivas hasta comprobar que todos podrían elegir los mismos principios. Por lo tanto, cada persona puede situarse en el punto arquimédico considerando las circunstancias en las cuales su elección de los principios de interacción social podrían converger con las elecciones de sus semejantes. Cada individuo puede

verificar que ha determinado correctamente esas circunstancias (de modo tal que su elección no refleje una deformación relacionada con su identidad particular) suponiendo que asume la identidad de cualquier otra persona seleccionada al azar y preguntándose si en las circunstancias consideradas esa otra persona hubiese hecho la misma elección, hubiese seleccionado los mismos principios.

Como el punto arquimédico exige que se ignore totalmente la propia identidad, podría suponerse que es imposible conectar a las personas reales, definidas por sus facultades, actitudes y preferencias, con el actor ideal. Pero esto podría constituir un error. Lo que conecta a cada persona con la elección ideal es la idea de que se trata de la elección hecha por un actor racional. Podemos, por decirlo así, trazar líneas de racionalidad desde las perspectivas reales en las cuales eligen las personas reales, hasta la perspectiva ideal brindada por el punto arquimédico. Esas líneas de racionalidad permiten determinar cómo progresivamente desaparecen las condiciones para conocer la propia identidad. En este sentido, el punto arquimédico es el punto que limita la identidad personal. Las elecciones convergen en ese límite porque desaparecieron los factores que provocan la divergencia. Cada persona se identifica con las elecciones a lo largo de la línea de racionalidad que se extiende desde su propia perspectiva, y por lo tanto, cada individuo se identifica, en el límite, con la elección hecha por el actor ideal.

Al elegir desde el punto arquimédico, la persona ideal elige como lo haría cada persona. Esta es la primera clave; la segunda clave que descorre el velo que oculta nuestra comprensión moral procede de una concepción clara de la persona individual considerada como actor, y por lo tanto como elector. Al criticar a Rawls nos esforzamos por resaltar lo esencial de esa concepción. El individuo real que debe identificarse con la elección ideal, tiene sus propias características particulares y definitorias: facultades, talentos, actitudes, preferencias. Obra considerando que tales características son dones y expresa su racionalidad en su esfuerzo por maximizar la satisfacción de sus preferencias utilizando sus facultades y otros rasgos de su carácter, en las circunstancias en las que fuera que se encuentre. En esto no participa ningún otro nivel de racionalidad. Y no hay otra

concepción de la persona implicada. Por conveniente que pueda ser para nuestros propósitos concebir a la persona de algún otro modo, tal concepción sería irrelevante para una teoría de la moral basada en la elección racional.

Los principios de interacción elegidos desde el punto arquimédico son pues principios para personas concebidas como actores racionales. Personas que deben atenerse a las condiciones de la acción racional y por lo tanto deben regular la interacción a la luz de la intención maximizadora de la utilidad de los actores racionales. Lo que hace que estos principios sean morales es el hecho de que regulen esa interacción sin referirse a las intenciones particulares, basadas en las características particulares, que tiene cada actor. Tales principios se atienen a la intención común a todos y lo hacen de un modo que impide que alguna expresión de tal intención avance sobre otra. El hecho de que la elección converja en el punto arquimédico logra este objetivo y asegura que se considere la intención común a todos y que no se pueda considerar la intención particular de ningún individuo. Lo primero suministra la condición positiva necesaria para que cada individuo se identifique con la elección; y el segundo punto proporciona la condición negativa que es igualmente necesaria, puesto que la inclusión de cualquier intención individual les impediría a las personas que no tengan esa intención identificarse con la elección. Por lo tanto, las líneas de racionalidad convergen en el punto que preserva toda individualización particular.

Pero hemos dicho que las intenciones individuales y las facultades y preferencias en las que se basan tales intenciones, quedan excluidas de la consideración moral, no porque sean arbitrarias, sino porque son parciales. El punto arquimédico neutraliza ese efecto. Aquí razonamos de un modo por completo diferente del de Rawls. Porque en su argumento, el efecto epistémico que tiene el hecho de ignorar la propia identidad adquiere una significación ontológica completamente incompatible con la idea de concebir a las personas como actores. Rawls llega así al punto de vista kantiano de que aparte del yo contingente, cada persona tiene un yo real, la persona moral, definida por su interés por la justicia y el bien, y ese yo real es el sujeto adecuado de toda elección moral.[30] Ese yo real, que es

lo único que escapa al nexo de la determinación social, se revela haciendo abstracción de todo conocimiento de los rasgos contingentes de la identidad individual. Pero nosotros negamos que exista un "yo real" semejante. La identidad de una persona es, en todos los sentidos, una cuestión contingente. Pero esa contingencia no es moralmente arbitraria, porque la moral sólo puede encontrarse (y sólo se encuentra) en la interacción de las personas reales individualizadas por sus facultades, sus actitudes y sus preferencias. En nuestro análisis tomamos en serio la individualidad de las personas.

4.2 El actor ideal razona partiendo de las condicines comunes a todos los individuos. Distingue la pregunta genética referente a cómo él llegó a ser un actor, de la pregunta normativa referente a cómo debería elegir atendiendo a su condición de actor y su respuesta a esta última pregunta es que elige racionalmente utilizar sus poderes a fin de satisfacer sus preferencias en la máxima extensión posible. Pero esto determina de dos maneras fundamentales y relacionadas entre sí, la elección que hace de los principios de interacción. En la medida en que esa persona puede elegir entre participar o no en la interacción, elige hacerlo sólo si espera beneficiarse de esa interacción. Y así, a fin de minimizar las ocasiones en las cuales debe participar de la interacción aunque no lo desee, y a fin de maximizar las oportunidades de participar de una interacción beneficiosa para sí, la persona elige una libertad de acción básica en relación con sus semejantes, una libertad que le permite promover sus propios intereses en la medida de lo posible, es decir, mientras los demás conserven la libertad de participar o no en la interacción con ella.

Que la sociedad sea "una empresa conjunta destinada al beneficio mutuo" es pues algo determinado por los términos de la elección hecha desde el punto arquimédico. Porque la elección del actor ideal es la elección de cada una de las personas. Si la persona ideal elige beneficiarse, debe pues elegir el beneficio mutuo; si elige la libertad para sí, debe pues elegir la libertad para todos. Al elegir los principios de interacción y las estructuras sociales que los encarnan, el actor ideal asegura pues, primero que cada persona comienza por tener una libertad fundamental para satisfacer sus preferencias y promover sus

intereses y, segundo, que las instituciones y prácticas sociales fueron creadas con el fin de brindar beneficio mutuo.

Estas dos exigencias —la libertad de acción y el carácter mutuo del beneficio— están vinculadas entre sí. Las personas que tienen la libertad de promover sus intereses cuanto les plazca siempre que esto no aleje a los demás de la interacción, sólo han de elegir la interacción cuando participar en ella les resulte beneficioso. Por otro lado, las personas que carecen de una libertad fundamental de acción pueden preferir entrar en interacción con las demás, a causa de las restricciones que limitan sus elecciones, aun cuando esa interacción no los beneficiara en el caso de que fueran libres. Por lo tanto, el actor ideal elige ser libre, de modo tal que pueda elegir sólo la interacción genuinamente beneficiosa y, al ser libre, naturalmente ha de elegir participar de la interacción únicamente de maneras que le resulten beneficiosas.

Al elegir la libertad de acción y el beneficio mutuo como producto de la interacción, el actor ideal sólo impone limitaciones muy débiles y difícilmente controvertidas a la cantidad de estructuras sociales aceptables. Aun así podemos preguntarnos si en algún momento de la historia humana esas limitaciones fueron respetadas por todos. Porque no basta con asegurar meramente que cada individuo se beneficie de todas las interacciones de las que participa con los demás consideradas en su conjunto, y comparadas con un hipotético estado de naturaleza en el cual el individuo viviría una existencia por completo solitaria. Antes bien es necesario que no se establezca ningún tipo de interacción que, considerada en todo su alcance, permita que algunas personas se beneficien imponiéndoles costos a otras. Si es posible impedir algún tipo de interacción, y si aquellas personas implicadas en ella la eligen libremente, esa interacción no puede brindarles a algunos la expectativa del beneficio general y brindarles a otros la expectativa del costo general.

La libertad elegida desde el punto de vista arquimédico no es por lo tanto la libertad de la interacción hobbesiana o anarquista, en la cual cada individuo puede perseguir su mayor beneficio sin preocuparse por los efectos que pueda causar esa búsqueda en las demás. Si la interacción ha de ser mutuamente

beneficiosa, debe prohibir que una persona imponga unilateralmente los costos a otra. Pueden existir —en realidad, en el mundo que conocemos deben existir— situaciones en las cuales la mera presencia de varias persona ocasiona desventajas mutuas y en un mundo superpoblado esas situaciones han de ser progresivamente más frecuentes. Pero no es nuestra intención enfrentar aquí al actor ideal con los problemas planteados por una elección de males. Suponiendo que exista la posibilidad del beneficio mutuo, el actor ideal debe elegir que se prohíba que alguien saque unilateralmente ventaja de algún otro.Nadie puede mejorar su propia situación mediante una interacción que empeore la posición de otro, si tomamos como punto de base (en relación con el cual determinamos qué es mejorar y qué es empeorar) la ausencia de la otra parte. Y así el actor ideal debe elegir la estipulación como uno de los principios de la interacción.

Al elegir la estipulación, el actor ideal no pone de manifiesto ningún interés altruista, ni siquiera imparcial, por sus semejantes. El actor ideal no tiene interés por refrenar sus impulsos de sacar ventaja de los demás, no siente deseos de dejar de mejorar su situación empeorando la de los demás, ya que ésta es sin dudas la forma más efectiva de maximizar su propia utilidad. Pero las condiciones de la elección arquimédica le impiden al actor ideal elegir principios de interacción que le permitan dar rienda suelta al interés exclusivo de maximizar su propia utilidad, cuando sólo puede satisfacerlo a expensas de los demás. Al elegir beneficiarse mediante la interacción, el actor ideal sólo puede elegir el beneficio mutuo. Y al elegir el beneficio mutuo, elige la prohibición de sacar ventaja, puesto que cualquier interacción que, considerada en todo su alcance pueda implicar el mejoramiento de la posición de una persona gracias al empeoramiento de la posición de otra, no puede brindar beneficio mutuo.

La estipulación es simultánemente la restricción más débil a las acciones de un individuo compatible con la exigencia de que, en la medida de lo posible, la interacción sea mutuamente beneficiosa y la restricción más fuerte a las acciones de un individuo compatible con su libertad de promover sus propios intereses. La estipulación define la posición inicial partiendo de

la cual se desarrollan la competencia de mercado y la cooperación, las formas de la interacción social que constituyen la intersección entre la libertad individual y el beneficio mutuo.

¿Qué alcance tiene la estipulación elegida por el actor ideal? ¿Es una restricción general a la acción? ¿O solamente una restricción a aquellas acciones particulares que se relacionan directamente con la competencia y la cooperación? El actor ideal sólo elige la interacción beneficiosa; podría por lo tanto parecer que la persona elige la estipulación sólo como una restricción limitada, cuya aceptación dispone a cada individuo a no sacar ventaja de sus semejantes sólo en el caso de que éstos participen en actividades conjuntas específicas. Pero esto sería un error. Porque el actor ideal reconoce que la disposición a esforzarse por obtener oportunidades de participar de la interacción cooperativa y de mercado es en sí misma ventajosa, atendiendo a los costos del estado de naturaleza y a los beneficios que ofrece la sociedad. De conformidad con la primera ley de la naturaleza de Hobbes, el actor ideal elige esforzarse por obtener la paz en la medida en que tenga la esperanza de obtenerla. Debe manifestar una voluntad permanente de comprometerse en la interacción social, de perseguir la ventaja mutua dónde y cuándo pueda encontrarla. Y así, al elegir los principios por los cuales ha de regularse la interacción, el actor ideal elige la estipulación como una restricción general, cuya aceptación muestra la disposición de cada persona a renunciar a su esfuerzo directo por maximizar su utilidad cuando todos pueden beneficiarse mediante la cooperación.

La sociedad, en su carácter de marco dentro del cual se desarrollan la competencia de mercado y la cooperación, es en sí misma una empresa cooperativa única. La estipulación es la restricción a la búsqueda que cada individuo emprende de su utilidad máxima y que hace posible semejante empresa. El actor ideal elige pues que cada individuo refrene sus impulsos de sacar ventaja de sus semejantes, aun cuando tal restricción a su conducta no se base en ningún acuerdo específico. Desde el punto arquimédico, la estipulación se considera no sólo como el factor determinante de la posición inicial de las transacciones particulares, sino también como el principio que, mutuamente aceptado por cualquier grupo de personas, suministra los ci-

mientos de una estructura de instituciones y prácticas mediante las cuales ese grupo de personas llega a constituir una sociedad.

El actor ideal, al disponerse a la interacción social, elige la estipulación, pero ¿hasta qué punto se dispone a abrazarla? La mayor parte de nosotros no está dispuesta a refrenar sus impulsos de sacar ventaja de los insectos; nos gusta la miel y hacemos lo que queremos con las abejas. ¿El actor ideal elige la estipulación como una restricción a toda interacción humana? ¿O es posible que el alcance de tal restricción sea limitado, de modo tal que restrinja las interacciones de los miembros de un género, una raza, una clase o una religión particular? Desde el punto arquimédico ¿es posible establecer una distinción entre *nosotros* y *ellos*, de modo que el actor ideal puede elegir que *nosotros* refrenemos el impulso de sacarnos ventajas entre nosotros y al mismo tiempo podamos caerles encima a *ellos* como zorros en el gallinero?

Si los intereses de ciertas personas o de ciertos grupos de personas fueran necesariamente opuestos, como en el caso de los dobus y las batatas o según la visión marxista de las clases, el actor ideal, aunque ignorara su identidad particular, reconocería pues que su esperanza de obtener la paz, su esperanza de beneficiarse de la interacción social, tiene una extensión limitada. Nadie puede identificarse con la elección de un principio que exija una restricción cuando no ha de obtenerse ningún beneficio. Si Marx tuviera razón, el actor ideal limitaría la estipulación a las actividades que se desarrollen dentro de una clase y dejaría que cada clase (o sus miembros) se esforzara libremente por beneficiarse a expensas de los demás.[31] Pero nosotros no aceptamos que necesariamente haya una oposición de intereses. El supuesto de los beneficios del mercado y de la interacción cooperativa se extiende a cada persona en interacción con todas las demás. El hecho de que haya ocasiones en las cuales las circunstancias inequitativas refutan este supuesto no puede afectar una elección hecha sin tener en cuenta ninguno de esos casos particulares. El actor ideal coincide con Hume en cuanto a que "la historia, la experiencia y la razón nos instruyen lo suficiente... y nuestro interés por la justicia va creciendo gradualmente a medida que vamos aceptando la dilatada utilidad de esa virtud".[32] Por lo tanto, elige la estipulación como una restricción a las interacción que mantiene toda la humanidad.

4.3 El actor ideal, después de haber establecido la estructura básica dentro de la cual ha de desarrollarse la interacción social, debe considerar los principios que regulen tal interacción. Primero, debe asegurar que él, y por lo tanto cada persona, disfrute de la mayor expectativa de utilidad compatible con la expectativa comparable de cada uno de los demás. El objeto de la elección arquimédica es no sólo el beneficio mutuo, sino también la optimidad.

Una vez que se da una libertad restringida por la estipulación, es decir que se logra que no haya uso del fraude ni de la fuerza en la interacción, surge el mercado. En condiciones de competencia perfecta, ese mercado produce un resultado óptimo que se alcanza mediante intercambios en los cuales cada individuo paga, y sólo paga, los costos de los beneficios que recibe. En la interacción de mercado no hay lugar para la injusticia; puesto que la estipulación excluye la injusticia al determinar los recursos que cada persona aporta al mercado, el resultado de éste es pues óptimo y justo. Y en él todos realizan libremente sus intercambios; el mercado respeta la libertad de cada individuo de elegir si participa o no en la interacción con los demás.

Por consiguiente, el actor ideal elige el mercado en condiciones de competencia perfecta. Ninguna otra elección es compatible con su defensa de la libertad restringida solamente por la estipulación y la optimidad. Puesto que en el mercado la persona paga los costos de los beneficios de los que disfruta, cualquier alternativa al mercado, que implique una distribución diferentes de los costos y los beneficios, necesariamente ha de incluir un desplazamiento de los costos mediante el cual algunos podrían mejorar su situación empeorando la de sus semejantes. Por lo tanto, cualquier alternativa restringiría la libertad de un modo que permitiría que alguno saque ventaja de otro.

Por lo tanto, el actor ideal no es un socialista. O, mejor dicho, no es un socialista ideal; no ha de elegir la propiedad social de los medios de producción en circunstancias que permitan que se dé el mercado perfectamente competitivo. No ha de elegir el socialismo, porque la restricción a la libertad individual exigida por la transferencia de los recursos de producción del individuo a la sociedad no podría justificarse ni mediante la estipulación ni mediante la demanda de optimidad. La solución a las violaciones de la estipulación que se producen al determinar los recursos disponibles iniciales de los individuos no es la colectivización de esas dotes, sino la redistribución justa de los recursos entre los individuos, de modo tal que pueda darse el mercado. La situación ideal para la elección arquimédica es la economía libre de mercado.

Pero recordemos que este argumento no es suficiente ni como defensa ni como condena de ningún arreglo social real. Porque tal argumento es perfectamente compatible con lo que ya dijimos referente a que las imperfecciones reales que se dan en los arreglos de mercado, que afectan no las dotes individuales sino la eficacia del intercambio voluntario al relacionar los costos con los beneficios, pueden justificar cierta forma de control social y hasta la propiedad social. Nuestro argumento es incompatible con la teoría socialista, del mismo modo en que es incompatible con la teoría utilitarista. Pero es posible, aunque improbable, que las diferentes teorías conduzcan a alguna práctica.

También hemos de observar que, al elegir el mercado, el actor ideal no se compromete con ninguna visión particular de herencia ni con ninguna visión del rédito general. Al relacionar la herencia o el patrimonio con el desinterés mutuo supuesto en la elección arquimédica, surgen problemas evidentes. Y nadie puede pretender como dotes previas al mercado aquel rédito que puede distribuirse sin que se afecte la eficiencia del mercado y que sólo es el producto de las relaciones de aquellos que participan en la interacción. Analizaremos la cuestión de la herencia y del rédito en el capítulo 9.

Pero no daremos por concluida esta subsección con estas calificaciones, por importantes que sean. Porque la conexión que existe entre la elección racional, la imparcialidad y la interacción del mercado, puesta de manifiesto por nuestro argumento, merece un análisis más profundo. El mercado, como lo vimos en el capítulo 4, crea una zona moralmente libre, pero es en sí mismo una creación racional y moral. Las condiciones para que se dé la interacción del mercado no siempre están presentes, pero su ausencia siempre provoca una pérdida para los actores racionales. Desde el punto arquimédico se elige el mercado como un ideal en sí mismo.

4.4 Establecidos la estipulación y el mercado, el actor ideal debe volcarse a la interacción en la que la presencia de las externalidades lo lleve a obtener un resultado subóptimo, si cada individuo es libre de promover sus propios intereses sin restricciones. ¿Cómo puede lograrse la optimidad, sin permitir que unos saquen ventaja de otros? Puesto que el actor ideal, al elegir su propio beneficio, debe elegir el beneficio mutuo, ¿cómo debe extenderse esa mutualidad a aquellas instituciones y prácticas que promueven la cooperación como la respuesta al fracaso del mercado?

La demanda de optimidad se satisface mediante la elección de una estrategia conjunta adecuada para la interacción, encarnada en la estructura social. Pero satisfacer la demanda de justicia, no basta para asegurar que la estructura social elegida garantice que, en cada tipo de interacción, nadie pueda esperar beneficiarse imponiéndole costos a algún otro. Porque la estructura misma puede favorecer diferentemente a algunas personas, promoviendo o reprimiendo ciertas potencialidades, exigiendo o prohibiendo ciertas actividades. Supongamos por ejemplo que dos personas tengan talentos, energías e intereses comparables a grandes rasgos, pero que una de ellas, que es varón, posea además una formación que lo ha preparado para desarrollar una carrera estimulante y socialmente valiosa, bien remunerada y altamente estimada, mientras que la otra, que es mujer, carece de esa educación y comprueba que como consecuencia de ello sólo puede dedicarse a una ocupación menos interesante, pobremente recompensada y de poca estima. Las retribuciones que reciban estos individuos pueden corresponder aproximadamente a las contribuciones que realmente hagan a la producción de los bienes. Cada persona paga por los beneficios recibidos. El hombre no obtiene un mayor beneficio imponiéndole costos a la mujer. Sin embargo las contribuciones que ambos hacen a la interacción social y por lo tanto sus retribuciones —que para asegurar el beneficio mutuo del principio al fin deben ser igualmente proporcionales a aquellas contribuciones—reflejan las contingencias sociales determinadas por una estructura social particular en la cual se impulsa a los varones a desarrollar aptitudes que se reprimen en las mujeres. Evidentemente hemos de considerar injusta esta estructura, no porque

permite que algunos individuos saquen ventajas de sus semejantes, porque en realidad no lo permite, ni tampoco porque no relacione los beneficios con las contribuciones, porque tampoco lo hace, sino que la consideramos injusta porque no relaciona los beneficios con las contribuciones que cada persona podría haber hecho si ambas gozaran de oportunidades similares y recibieran un apoyo semejante.

A fin de superar esta dificultad debemos suponer que el actor ideal considera no sólo el funcionamiento interno de cada una de las estructuras sociales entre las cuales debe elegir sino también los funcionamientos comparativos de las diferentes estructuras. La estructura elegida no sólo debe asegurar que ningun individuo se beneficie realmente a expensas de otro sino también que ningún individuo se beneficie de manera diferente gracias a esa particular estructura social. El interés del actor ideal, lo repetimos, no es en absoluto altruista ni imparcial. A él sólo le importa asegurarse que la interacción social lo beneficie, no meramente en una situación aislada sino del principio al fin. Pero esto le exige elegir principios de interacción que aseguren el beneficio mutuo pleno en cada tipo de interacción, considerados en relación con las alternativas factibles.

Por consiguiente, los principios elegidos desde el punto arquimédico deben lograr que la porción que cada persona espera obtener de los frutos de la interacción social se relacionen, no con lo que esa persona realmente aporta, puesto que su contribución real puede reflejar los permisos y las prohibiciones contingentes que existen en cualquier estructura social, sino con la contribución que la persona podría hacer en una estructura social más favorable para el desarrollo de sus aptitudes y sus rasgos de carácter y para la satisfacción de sus preferencias, siempre que tal estructura fuera una alternativa factible que satisficiera las otras exigencias de la elección arquimédica. (Una estructura en la que yo soy el emperador y los demás son mis humildes sirvientes, aunque sea favorable a mi megalomanía queda descalificada porque ignora abiertamente la estipulación.) A fin de medir la contribución potencial de un individuo en esa admisible estructura social que le es más favorable, la relacionamos, siguiendo el argumento expuesto en 5.4.3 con la pretensión que ese individuo podría tener sobre el superávit

obtenido mediante la interacción social. Si llamamos a esa pretensión su beneficio social máximo, podemos decir que, al elegir una estructura social, el actor ideal debe lograr que en la medida de lo posible todos los individuos puedan esperar proporciones iguales de sus beneficios sociales máximos. Así, el actor ideal relaciona la elección de una estructura social con la idea de la sociedad como una única empresa cooperativa.

Para que no se piense que la idea de la contribución potencial es algo completamente indeterminado, puesto que las diferentes condiciones sociales e históricas modelan a los seres humanos de maneras incomparablemente distintas, deberíamos señalar que la comparación entre estructuras sociales exigida para determinar la contribución potencial y el beneficio social máximo debe limitarse a las estructuras que constituyen alternativas factibles. El actor ideal no pasa revista a las sociedades pasadas, presentes y futuras, del mundo real y de todos los mundos posibles para encontrar la que mejor se ajusta a cada persona individual. Antes bien compara las circunstancias históricas y sociales reales en las cuales un individuo se encuentra con las alternativas disponibles en el tiempo y en el lugar que le corresponden, con la tecnología de que dispone y el horizonte conceptual que existe. Esos límites son vagos, pero hacen posible la comparación significativa exigida.

El actor ideal elige que nadie saque ventaja de él en la interacción. Por lo tanto, elige principios que determinen una estructura social en la cual su expectativa del beneficio que pueda brindarle la cooperación en relación con el beneficio máximo que podría recibir en alguna estructura social factible no sea menor que la recibida por sus semejantes, o si la equidad es subóptima, que no sea menor que la que deba recibir alguna persona sea cual fuere la estructura elegida. Si la persona fuera consciente de su identidad, y estuviera en posición de elegir entre diferentes principios de interacción, podría hacer y por supuesto haría una elección más ventajosa que reflejara sus habilidades y preferencias particulares, pero desde el punto arquimédico no puede hacerlo. Dada su ignorancia, la persona sólo puede elegir maximizar el beneficio relativo mínimo que puede brindarle la interacción social. Y así, desde el punto arquimédico, elige el principio de concesión relativa minimáxima, puesto que como mostramos en 5.4.3, es equivalente al beneficio relativo maximínimo. La estipulación, el mercado y la cooperación basada en la concesión relativa minimáxima expresan juntos los intereses fundamentales del actor ideal, la libertad de la acción individual y el beneficio mutuo obtenido mediante la interacción, y cómo esos intereses se extienden y se aplican a la sociedad concebida ésta como una estructura general destinada a la cooperación.

4.5 Los principios de interacción social elegidos desde el punto arquimédico son los mismos principios del acuerdo racional en los que se basa la cooperación. Esto no debería sorprendernos. Porque si cada persona debe identificarse con la elección hecha por el actor ideal, su elección debe ser el resultado, no de una decisión individual tomada en situación de incertidumbre, sino de una negociación o un contrato que todos los individuos puedan aceptar. El actor ideal no puede desbaratar las identidades de las personas (él podría ser cualquiera de ellas) y convertirla en una única identidad, ni desintegrar todas las utilidades y transformarlas en una sola para luego tratar de maximizarla. Debe mantener separadas todas las identidades y todas las utilidades y por lo tanto debe elegir como si negociara con cada uno de sus semejantes. Este es el rasgo principal de la elección arquimédica.

El punto de Arquímedes puede parecer un artificio peculiar, presentado únicamente para asegurar la imparcialidad de la elección. Y podríamos preguntarnos qué es para nosotros el punto arquimédico. ¿Por qué preocuparnos con principios elegidos en una condición limitadora de ignorancia de la propia identidad impuesta por ese punto? Pero esa ignorancia es conveniente para seleccionar aquellos principios encarnados en las instituciones y prácticas que constituyen nuestra matriz social. Al reflexionar sobre los procesos de socialización que influyen profundamente en el momento de determinar los individuos particulares que somos, comprendemos que para evaluarlos y para evaluar los principios subyacentes en tales procesos debemos encontrar un punto con el cual cada uno de nosotros pueda identificarse y que sin embargo, sea independiente de nuestras características particulares. Ese punto es exactamente el de la elección arquimédica. Al identificarnos con

las elecciones del actor ideal justificamos ante nosotros mismos los procesos sociales mediante los cuales llegamos a ser quienes somos. Participamos de la elección de aquello que literalmente no podemos elegir de un modo que le dé significación a la idea de nuestra autonomía racional. Una persona autónoma es aquella que se identifica conscientemente con los procesos que la modelaron como actor racional.

Lejos de presentar una perspectiva peculiar y artificial para la elección, el punto arquimédico ofrece el punto de vista familiar del acuerdo. Una decisión tomada en completa ignorancia de la propia identidad coincide con un acuerdo hecho entre personas individualizadas de todos los modos posibles. Las líneas de racionalidad que parten de cada individuo y convergen en el actor ideal son las mismas líneas que parten de los individuos y convergen en una negociación racional. Por lo tanto, el punto arquimédico es la posición que ocupa cada uno de nosotros cuando participamos voluntaria e igualmente de un acuerdo con nuestros semejantes. La elección arquimédica da origen al contrato social.

No obstante, la elección arquimédica no origina por sí misma el acatamiento del contrato social. Aunque cada persona se identifique con la elección de los principios sociales hechos por el actor ideal, no por ello se compromete racionalmente a acatar personalmente esos principios. La racionalidad del acatamiento, como sostuvimos en el capítulo 6, depende de las características y las circunstancias de los individuos que participan de la interacción. Y no puede ser determinada por el punto arquimédico.

El actor ideal elige, no el acatamiento, sino aquellos procesos de socialización que promueven las circunstancias en las cuales resulta racional acatar estrechamente. Atendiendo a los beneficios que cada individuo puede esperar obtener de la cooperación, cada uno tiene razones para preferir que todos sean suficientemente transparentes para que la disposición a acatar la estipulación y el principio de concesión relativa minimáxima sea en sí misma maximizadora de la utilidad. Y cada individuo tiene razones para preferir que todos se sientan comprometidos afectivamente con el acatamiento, de modo tal que los sentimientos familiares de respeto y ofensa, de autoestima y culpa, se vinculen apropiadamente con la conducta justa e injusta de

los demás y de uno mismo. Aunque nuestra preocupación principal se refiera a los principios que uno elegiría desde el punto arquimédico, no debemos olvidar la importancia de la elección de los afectos, en la medida en que la socialización pueda influir en la forma que éstos adquieran.

La teoría moral ofrece un punto de Arquímedes para analizar la interacción humana. La teoría de la elección racional permite hacer un análisis desde el punto de vista de cada individuo que participa de la interacción. Al mostrar que hay armonía entre estos dos análisis hemos puesto de manifiesto que la demanda moral de imparcialidad puede ajustarse a la demanda racional de maximización de la utilidad individual cumpliendo las condiciones de interacción acordadas entre personas igualmente racionales. Esto completa nuestra exposición teórica; nos queda por presentar un esbozo de cómo una moral basada en el acuerdo puede aplicarse a la evaluación de las prácticas sociales y una discusión sobre la viabilidad de la vida social basada en una moral contractual.

Notas

1 J. Rawls, *A Theory of Justice*, Cambridge, Mass., 1971, pág. 584.

2 El carácter kantiano de la teoría de Rawls se hace cada vez más evidente en estos trabajos más recientes. Véase "Kantian Constructivism in Moral Theory", pág. 515.

3 Véase *Republic* 620d-621b.

4 Sobre ese ideal véase Kant *Groundwork of the Metaphisic of Morals*, págs. 425-426.

5 Rawls, *A Theory of Justice*, pág. 256.

6 *Ibid.*, págs. 255-256, relaciona el yo nouménico con la idea del actor racional.

7 Los defensores de las teorías del "observador ideal" ignoran este punto y suponen que la observación desencadena la reación. Véase, por ejemplo, el enfoque de R. B. Brandt, *Ethical Theory*, Englewood Cliffs, NJ, 1959, págs. 173-174.

8 Véase el análisis de la simpatía que hace Hume en *Treatise*, ii, i, xi, págs. 316-320.

9 Véase Kant, *Groundwork of the Metaphysic of Morals*, pág. 398.

10 Hobbes, *De Cive*, XV, 7, pág. 294.

11 Véase J. C. Harsanyi, "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking", en *Essays on Ethics*, págs. 3-5. Analizo exhaustivamente la teoría de Harsanyi en "On the Refutation of Utilitarianism", en H. B. Miller y W. H. Williams (comps.), *The Limits of Utilitarianism*, Minneapolis, 1982, págs. 144-163.

12 Harsanyi, "Rule Utilitarianism and Decision Theory", *Erkenntnis* 11, 1977, pág. 28.

13 Véase Mill, cap. 2, párrafo 6.

14 Véase Harsanyi, *Rational Behavior in Games and Social Situations*, Cambridge, 1977, págs. 51-55, en donde el autor defiende la idea de dejar de lado esas preferencias al desarrollar la explicación apelando al principio de la soberanía del consumidor.

15 Rawls, *op. cit.*, pág. 27.

16 Véase *ibid.*, págs. 12, 136-142. Sobre un análisis parcialmente superpuesto de la teoría de Rawls véase mi "Justice and Natural Endowment: Toward a Critique of Rawls' Ideological Framework", *Social Theory and Practice* 3, 1974, págs. 3-26.

17 Rawls, *op. cit.*,pág. 62.

18 *Ibid.*, pág. 83.

19 *Ibid.*, págs. 62, 92.

20 Estos argumentos se encuentran primariamente en *A Theory of Justice*, págs. 150-161, 175-183.

21 Véase *ibid.*, págs. 167-175.

22 *Ibid.*, págs. 101-102.

23 Rawls, "The Basic Structure as Subject", *American Philosophical Quarterly* 14 (1977), pág.62.

24 Rawls, *A Theory of Justice*, pág. 4.

25 *Ibid.*, pág. 101.

26 Rawls, "The Basic Structure as Subject", pág. 162; está invertido el orden de las dos partes de la cita.

27 Rawls, *A Theory of Justice*, págs. 74-75.

28 Véase Rawls, "The Basic Structure as Subject", pág. 162.

29 En "Kantian Constructivism in Moral Theory", Rawls sugiere una respuesta por completo diferente a la pregunta sobre por qué los haberes naturales son moralmente arbitrarios. Allí el autor presenta el problema de la elección arquimédica como la adopción de "principios que sirvan como concepción pública efectiva de la justicia para una sociedad bien ordenada", en la cual los ciudadanos se consideren "personas morales...caracterizadas por dos fuerzas morales y dos intereses correspondientes del más alto orden por desarrollar y ejercer esas fuerzas. La primera fuerza es la capacidad de tener un sentido efectivo de la justicia... La segunda fuerza moral es la capacidad de formar, corregir y perseguir racionalmente una concepción del bien. Coherentemente con estas fuerzas morales, las personas están movidas por dos intereses del orden más alto...que son reguladores y efectivos en grado sumo. Esto implica que, sean cuales fueren las circunstancias relevantes para su satisfacción, esos intereses gobiernan la deliberación y la conducta." (págs. 524-525). En esta perspectiva, el carácter ideal del actor situado en el punto arquimédico estriba en la preocupación exclusiva que siente por los intereses del orden más alto, al considerar cómo han de distribuirse los frutos de la interacción. Pero en esta perspectiva los principios se eligen con el fin de regular la interacción, no de maximizadores racionales de la utilidad, para quienes la moral es una cuestión de acuerdos, sino de personas que ya eran previamente morales. La preocupación por la moral es algo ya dado y el punto arquimédico debe pues expresar esa preocupación. Sencillamente éste no es nuestro problema.

30 Véase Rawls, *A Theory of Justice*, págs. 255-256.

31 Véase J. Elster, *Ulysses and the Sirens: Studies in rationality and irrationality*, Cambridge, 1979, pág.97, sobre una esclarecedora comparación de Hobbes y Marx que resulta relevante en este aspecto.

32 Hume, *Enquiry*,iii, i, pág. 192.

IX

Personas, pueblos, generaciones

John Locke escribió que "un hobbesiano no ha de admitir fácilmente que haya muchos simples deberes morales".[1] Aunque nuestra manera de entender la motivación y la racionalidad humanas es menos restrictiva que la de Hobbes, la teoría moral que hemos desarrollado es claramente hobbesiana en cuanto a la importancia que le da a la restricción mutua. Una perspectiva más amplia nos permitió establecer los derechos individuales de un carácter reconociblemente lockeano, así como los principios de justicia. Sin embargo podemos aceptar que las restricciones morales que surgen de lo que, en su sentido más amplio, son las condiciones de ventaja mutua, no corresponden en todos los aspectos a los "simples deberes" de la moral convencional. Los animales, los no nacidos, las personas congénitamente incapacitadas y defectuosas, están más allá de la esfera de una moral vinculada con la mutualidad. La disposición a acatar las restricciones morales, sin las cuales las relaciones morales son objeto del desdén del necio, sólo puede defenderse racionalmente dentro del alcance del beneficio esperado.

Al mostrar que la concesión relativa minimáxima es un principio de justicia y también de elección racional, damos a la moral el fundamento seguro que se exige ante todo en el desafío que Glaucón y Adimanto le presentan a Sócrates.[2] Pero (para tomar prestada una paradoja publicitaria) nuestro argumento nos da todo lo que siempre quisimos de una moral...y menos.[3] Nos da todo, al proporcionarnos un fundamento racional que no somete a un tira y afloja a nuestra concepción de la racionalidad; y nos da menos al proporcionarnos ese fundamento sólo dentro de la esfera del beneficio mutuo.

En este capítulo consideraremos qué implica exactamente nuestra teoría moral en la tarea de evaluar algunas de las relaciones medulares de la moral social. Hemos visto ya que nuestra teoría moral da lugar a normas muy diferentes para evaluar la distribución de los bienes públicos y privados. Consideraremos ahora qué implicaciones tiene nuestro argumento para la distribución de la renta bruta y mostraremos que nuestra posición se sitúa entre el sencillo individualismo de Robert Nozick y el colectivismo implícito de John Rawls. Examinaremos además las implicaciones que tienen las diferencias de adquisición inicial en las relaciones entre las personas y las naciones. Mostraremos cómo, con la ayuda de supuestos empíricos muy simples y plausibles, podemos ordenar parcialmente diferentes estilos de vida atendiendo a sus niveles de progreso y justificar que un estilo más avanzado reemplace a otro menos avanzado. Mostraremos cómo las relaciones morales que mantienen los miembros adultos normales de la sociedad pueden extenderse hasta abarcar a los jóvenes, como futuros contribuyentes, y a los ancianos, como antiguos contribuyentes, y lo haremos dándoles un explicación racional contractual a las prácticas de herencia e inversión, explicación que parece evitar las consecuencias no intuitivas del utilitarismo y del enfoque de Rawls.

Veremos que no será sencillo probar que nuestra teoría está en armonía (o no lo está) con los "deberes supuestamente simples" de la moral convencional. Al confiar en la teoría antes que en la intuición, defenderíamos el enfoque de las relaciones sociales esbozado en este capítulo pasando por alto las modas intelectuales del momento.[4] Si el lector siente la tentación de objetar alguna parte de este enfoque, alegando que se violan sus propias intuiciones morales, debería preguntarse qué peso puede tener tal objeción si la moral ha de obrar dentro de la esfera de la elección racional. Queremos destacar la diferencia radical que existe entre nuestro enfoque, en el cual no preguntamos qué visión de las relaciones sociales podría ser aceptada *ex ante* por individuos preocupados por maximizar sus utilidades, y el enfoque de los "coherentistas" morales y los defensores del "equilibrio reflexivo", quienes le dan importancia inicial a nuestros juicios morales considerados.[5]

No estamos diciendo que el único modo de refutar las afirmaciones que hacemos en este capítulo sea presentar una teoría moral superior. Porque podemos haber malinterpretado las implicaciones de nuestra propia posición. Al tratar de aplicar una teoría moral a circunstancias mucho más complejas que

aquellas imaginadas en el marco idealizado empleado para elaborarla, se presentan enormes dificultades. De modo tal que en este capítulo adoptamos un espíritu explorador, y distinguimos aquellas que parecen ser inequidades justificables de las injustificables, distinguimos entre los progresos y los retrocesos humanos, pero estamos dispuestos a admitir que aun cuando nuestra teoría de una moral basada en el acuerdo fuera plenamente aceptable, gran parte de lo que afirmamos aquí sólo tiene un carácter tentativo y ha de prestarse a controversias.

2 En nuestra teoría la igualdad no ocupa un lugar fundamental. Al sostener que las instituciones y prácticas sociales están justificadas por la apelación a un hipotético acuerdo *ex ante* establecido entre personas racionales y plenamente informadas, nos hemos referido a la igual racionalidad de los negociadores, a fin de mostrar que el acuerdo convenido entre ellos satisface la norma moral de imparcialidad. Pero las prácticas imparciales respetan a las personas tales como éstas son, respeta tanto las igualdades como las desigualdades que haya entre ellas. Puesto que el mercado perfectamente competitivo suministra un marco para la interacción que deja a cada individuo disfrutar de la plena libertad de que goza un Robinson Crusoe, no debe sorprendernos que las desigualdades naturales que puedan manifestar los diferentes Robinson Crusoe reaparezcan como desigualdades del mercado. Puesto que nadie está en posición de sacar ventaja de ningún otro, las desigualdades del mercado puro no revelan una injusticia mayor que la que puede revelarse entre los individuos aislados que no están en interacción. Aquellos capaces de aportar más servicios intermediarios al mercado pueden esperar disfrutar de una porción preferida de los bienes privados, tales como alimentos, vestidos, vivienda, transporte privado y recreación que proporciona el mercado.

Las sociedades de mercado puro son con frecuencia objeto de críticas a causa de las desigualdades que supuestamente fomentan. Pero, si bien la desigualdad del mercado es una base legítima para que se lo critique, la verdadera fuente de sus yerros es la presencia de las externalidades que permiten la aparición de los jugadores líberos y los parásitos. El mercado suministra bienes públicos de un modo que no es ni eficiente ni

justo. En su adecuada extensión, el mercado es en realidad una fuerza igualadora de los asuntos humanos que elimina las violaciones a la estipulación lockeana que caracterizan a las sociedades de esclavos, siervos y castas. El efecto que produce subordinar el mercado a alguna forma de control político es la manera típica de reintroducir las violaciones a la estipulación y de atropellar sin miramiento alguno la justicia, como lo mostraron bien las llamadas sociedades socialistas del Este. En lugar de suprimir el mercado, lo que hace falta es complementarlo con instituciones y prácticas cooperativas que aseguren la oferta óptima de bienes públicos.

La cooperación da lugar a una distribución mucho más igualitaria de los bienes producidos que la que brinda el mercado. Porque, como lo mostramos en 5.4.2-3, al producir el superávit cooperativo los diferentes individuos hacen contribuciones equivalentes. Puesto que ese superávit constituye un bien único transferible, los cooperadores contribuyen igualmente a producirlo y, por lo tanto, tienen derecho a compartirlo igualmente. Ahora bien, podría suponerse que en el caso de los bienes puramente públicos, definidos por su carácter no excluyente, la distribución igual queda asegurada.[6] Si una sola persona proveyera la defensa nacional y el aire no contaminado, éstos serían bienes destinados a todos. Y si todos deben recibir su porción de los bienes, los derechos de cada uno son pues irrelevantes. Pero puesto que el carácter excluyente y el carácter no excluyente son límites ideales, la mayor parte de los bienes que constituyen el superávit cooperativo pueden ser suministrados diferencialmente; las instituciones educativas públicas no necesariamente tienen que abrir sus puertas a todos. Y, en todo caso, lo que según nuestro argumento debe distribuirse equitativamente es el superávit neto; aun cuando el bien de la defensa nacional o del aire no contaminado deba ser suministrado equitativamente a todos, no es necesario que todos carguen con los costos equitativamente.

¿Acaso estamos comprometiéndonos con la idea de que los bienes puramente públicos deben suministrarse sobre la base de contribuciones iguales de todos los miembros de la sociedad? Si la defensa nacional es un bien que se les provee a todos por igual, ¿debe estar sostenida por contribuciones iguales de todos?

Si esto fuera lo que exige el principio de concesión relativa minimáxima, seguramente serían irracionales y además injustos no sólo el sistema existente de impuestos progresivos sino también la alternativa de tarifa fija, ampliamente defendida por economistas tales como Milton Friedman.[7] En la medida en que los bienes producidos de manera cooperativa o pública se dividan equitativamente entre los miembros de la sociedad, los impuestos también deben ser iguales, de modo tal que cada individuo reciba una porción equivalente del excedente cooperativo neto.

Este argumento es plausible sólo aparentemente, pues se basa en ignorar que la concesión relativa minimáxima se aplica a la distribución del superávit cooperativo considerado no como un conjunto de bienes sino como un conjunto de utilidades individuales. Consideremos una negociación racional sobre la división de los costos de producir un bien puramente público. En el resultado no cooperativo, que sirve como posición inicial, el bien no se produce. Cada persona pretende el nivel máximo del bien que pueda producirse sin provocarle costos y que deje a cada uno de los demás individuos por lo menos tan bien situado como lo estaría si el bien no se produjera en absoluto. De modo que el nivel real en el que las personas acuerdan producir el bien y la división real de los costos entre ellas, se determinan aplicando la concesión relativa minimáxima o, de manera equivalente, el beneficio relativo maximínimo. Para determinar cada nivel y cada división de costos posibles, la ganancia neta de cada individuo se mide como una proporción de la ganancia que tal individuo obtendría de su pretensión y se seleccionan el nivel y la distribución maximizadores de la ganancia neta proporcional mínima.

En general, si dos personas pagan equitativamente, en dinero o en recursos, por un bien puramente público, la persona con menores recursos perdería pues una proporción mayor de la ganancia que obtendría si recibiera el bien sin costo alguno. Por consiguiente, esa persona hace una concesión relativa mayor que la otra. Pero normalmente, el principio de concesión relativa minimáxima iguala las concesiones relativas. Por lo tanto, no exige que las personas paguen impuestos iguales, sino que antes bien exige pagos mayores a aquellos más beneficiados. De modo

que es plausible suponer que los impuestos de tarifa fija proporcionan beneficios relativos iguales. Por supuesto, la tarifa misma debe determinarse de modo tal que brinde una oferta óptima de bienes producidos públicamente.

Los bienes privados se proveen separadamente a precios fijos. Los bienes públicos se proveen en un paquete único a precios que varían según los recursos de los compradores. La cooperación resulta ser más equitativa en sus efectos que el mercado. En una sociedad regida por el principio de concesión relativa minimáxima, la protección policial y judicial, la defensa nacional, la tierra, el aire y el agua libres de contaminación son bienes que se les suministran a todos sobre la base de impuestos proporcionales. Las maneras intuitivamente diferentes que tenemos de considerar las desigualdades de la distribución de los bienes públicos y privados reflejan no la incoherencia de nuestras perspectivas, sino una diferencia profunda en el modo en que funcionan las instituciones sociales que satisfacen una norma única de racionalidad e imparcialidad.

3.1 Al analizar la interacción del mercado omitimos una desigualdad de la retribución que no está determinada por el costo que implica proveer los servicios intermediarios. Como observamos en 4.3.2, en el mercado las personas pueden recibir un rédito. El rédito está determinado por la escasez; es el premio que obtienen ciertos servicios intermediarios, por encima del costo completo de proveerlos, porque no hay una forma alternativa de satisfacer la demanda.

Existe una demanda de hockey expresada en la venta de entradas a los partidos de hockey. Esa demanda queda satisfecha por las actividades de los jugadores, quienes, por consiguiente, pueden obtener por sus servicios un pago que se ajuste a esa demanda y supere los costos. Pero existe una demanda adicional por la habilidad única de Wayne Gretzky, expresada por la ocupación del estadio, antes a medias vacío, cada vez que Gretzky y los Edmonton Oilers se presentan en la ciudad. No se dispone de ningún sustituto que pueda satisfacer esta demanda. Por lo tanto Gretzky está en posición de obtener por sus servicios un pago que supere sus costos, incluyendo el costo de oportunidad, que le provoca suministrar tales servicios; está en posición de obtener un rédito. Seguramente, Gretzky estaría

dispuesto a jugar por una paga menor de la que recibe; la diferencia entre la cantidad mínima de dinero que lo induciría a jugar tan bien como lo hace y su remuneración real es pues su rédito.

Supongamos que en la práctica la renta bruta pudiera determinarse y se le impusiera un impuesto del 100 por ciento. La distribución del superávit representado por el rédito no afecta la optimidad del resultado. Aun un impuesto confiscatorio que provocara una redistribución completa del rédito sería meramente redistributivo en sus efectos. No afectaría ni podría afectar la oferta de los servicios y por lo tanto tampoco la producción de los bienes. El impuesto sobre la remuneración del mercado, generalmente afecta los funcionamientos del mercado de una manera adversa, pues cambia los términos en los cuales se proveen los servicios intermediarios. Pero un impuesto sobre el rédito, por alto que pueda ser, no conduciría a ninguna forma de ineficiencia.

El impuesto al rédito no sólo no puede afectar la optimidad, sino que tampoco puede afecta la libertad natural, asegurada por el mercado. Cada persona seguiría siendo capaz de dirigir sin interferencia sus habilidades al servicio de sus preferencias. El hecho de saber que no podrá obtener la renta bruta no incita a ninguna persona racional a alterar su interacción con las demás. Si Wayne Gretzky elige racionalmente jugar al hockey, el costo de oportunidad que tiene para él el hecho de jugar queda determinado por el beneficio neto que podría esperar obtener de la mejor alternativa disponible diferente de jugar al hockey. Si, como resultado del impuesto anticipado sobre la remuneración que le corresponde por jugar al hockey, Gretzky decidiera elegir otra carrera, ese impuesto necesariamente afectaría su costo de oportunidad y por lo tanto no se estaría gravando únicamente su rédito. Un impuesto sobre el rédito no puede afectar el orden de preferencias de los cursos alternativos de acción y por lo tanto no puede afectar la libertad del individuo.

Podría suponerse que la estipulación, al reconocerle a cada persona el derecho a sus dotes básicas —sus habilidades naturales— le reconoce el derecho al rédito que deriva de esas dotes. Wayne Gretzky tiene derecho a su pericia excepcional; puede utilizarla como mejor le plazca; ¿no tiene entonces el derecho al

rédito que esas habilidades le proporcionan? Pero la distribución del rédito no afecta su derecho a utilizar sus habilidades de jugador de hockey como mejor le plazca; sencillamente hemos mostrado que un impuesto confiscatorio sobre el rédito no afectaría y no debería afectar su disposición a jugar al hockey. El derecho de cada persona a su dote básica es un derecho al uso exclusivo de esas dotes en el mercado y en la interacción cooperativa. Pero la distribución del superávit representado por el rédito no afecta la interacción del mercado; si se le confisca el rédito, una persona no pierde el derecho al uso exclusivo de sus habilidades en la interacción del mercado. Pero, ¿qué ocurre con el derecho que cada persona tiene a sus propias habilidades en la interacción cooperativa?

Puede concebirse la sociedad como una empresa cooperativa única. El beneficio representado por la renta bruta es parte del superávit producido por esa empresa, porque éste surge sólo de la interacción social. Pero luego ese beneficio debe distribuirse entre los miembros de la sociedad en los términos establecidos por la concesión relativa minimáxima. Cada persona, en su carácter de contribuyente de la interacción social, comparte la producción del beneficio representado por la renta bruta. El talento de Wayne Gretzky obtiene un rédito porque es escaso, pero esa escasez no es una característica inherente a sus habilidades, sino que es una función de las condiciones de la oferta y por lo tanto de la relación existente entre su talento y el talento de los demás, y también una función de las condiciones de la demanda, y por lo tanto de la relación existente entre su talento y el interés de los demás por asistir a los partidos de hockey.

El derecho a usar las propias habilidades como mejor encajen en la cooperación con los demás no implica sin embargo un derecho a recibir la renta bruta, sino que sólo implica un derecho a compartir con los demás el superávit representando por esa renta. En la medida en que ese superávit adquiera la forma de un bien único transferible, cada miembro de la sociedad tiene derecho a una porción igual del mismo. Si aplicamos el principio de concesión relativa minimáxima a los términos generales de la cooperación quedan excluidas de la renta bruta todas las desigualdades. Pero debemos tener cuidado de aplicar

este resultado general igualitario. Porque no es fácil distinguir el rédito del pago por los costos de la oferta. En muchas actividades, la disposición que muestran las personas a mejorar al máximo su rendimiento puede relacionarse con las recompensas que reciben aquellos que tienen más éxito. El alto nivel de remuneración que reciben unos pocos puede pues ser no una renta bruta, sino antes bien un "premio gordo". Así como los premios de la lotería implican una desigualdad sustancial al ofrecer unos pocos premios mayores para inducir a muchos a participar, en muchas actividades humanas la desigualdad sustancial de las recompensas puede llegar a ser la condición necesaria para inducir al esfuerzo a los más. Es posible que Wayne Gretzky reciba no sólo la renta bruta sino también el premio gordo necesario para alentar a todos los jugadores de hockey a perfeccionarse. El superávit cooperativo no incluye pagos que cubran los costos de la oferta ni los pagos desiguales destinados a inducir al esfuerzo para cubrir tales costos. En la práctica, la distribución equitativa de la renta bruta puede producir muy poco efecto en las desigualdades de la riqueza.

3.2 Los análisis recientes sobre la renta bruta hechos por algunos teóricos de la moral llegaron a conclusiones por completo diferentes de la nuestra. Los lectores familiarizados con *Anarchy, State, and Utopia* de Robert Nozick sin duda habrán reconocido en nuestra discusión sobre Wayne Gretzky el intencionado paralelo con el tratamiento que hace Nozick de Wilt Chamberlain (los norteamericanos hablan de los jugadores de básquetbol y los canadienses de los jugadores de hockey).[8] A Nozick le interesa mostrar cómo "la libertad perturba las estructuras".

Empiece con cualquier distribución de bienes conforme con su concepción preferida.

> Ahora supongamos que Wilt Chamberlain es objeto de una gran demanda por parte de los equipos de básquetbol porque atrae mucho a los espectadores... De modo que Chamberlain firma con un equipo un contrato semejante a éste: en cada partido jugado en el país, Chamberlain recibe veinticinco centavos de dólar del precio de cada entrada... Comienza la temporada y la gente asiste encantada a verlo jugar; al comprar los billetes cada espectador deposita en una caja especial que lleva el nombre del jugador estrella veinticinco centavos del precio total... Supongamos que en una temporada un millón de personas asisten a los partidos de Chamberlain y que por lo tanto él recibe $250.000, una suma mucho mayor que los ingresos promedio de los jugadores de básquetbol y aun mayor que la de ningún otro. [Recordemos que cuando Nozick escribió esto, en los lejanos comienzos de la década de 1970, las cifras que se pagaban en los deportes profesionales eran por supuesto mucho menores que las actuales.] ¿Tiene derecho Chamberlain a obtener esos ingresos? Esta nueva distribución, ¿es injusta? Y si lo es, ¿por que?[9]

Por supuesto, Nozick quiere mostrar que la nueva distribución no es injusta. Cada persona que paga para ver jugar a Chamberlain elige voluntariamente hacerlo; cada persona se beneficia y Chamberlain también se beneficia. Cada persona tiene derecho al dinero que posee (es parte de una distribución de bienes que satisface las exigencias de la justicia); ¿no debería Chamberlain tener derecho al resultado de la transferencia que cada espectador hace voluntariamente de aquello que inicialmente le pertenecía?

¿Cuál es la naturaleza del pago que recibe Chamberlain? ¿La perspectiva de obtener $250.000 es lo único que lo induce a jugar otra temporada de básquetbol? Si los términos del contrato que él propuso fueran rechazados, ¿preferiría el jugador retirarse? Si así fuera, el pago que recibiría cubriría únicamente el costo de su oferta (en este caso, primariamente el costo de oportunidad) por el servicio que brinda al jugar básquetbol. El contrato representa un intercambio de mercado estándar y tenemos que convenir con Nozick en que interferir en él constituiría una negación de la libertad.

¿O Chamberlain estaría dispuesto a jugar aun cuando sólo recibiera, digamos, $100.000? Si así fuera, la porción excedente del pago de $100.000 constituiría un rédito e interferir en él no afectaría la decisión de Chamberlain de jugar. Nozick supone que Chamberlain tiene derecho a cualquier rédito que pueda obtener mediante el intercambio voluntario (restringido por una forma de la estipulación lockeana). Pero nosotros sostenemos que los términos de la cooperación social, establecidos por el principio de concesión relativa minimáxima, descartan la posibilidad de un contrato que incluya la renta bruta. Ciertamente el principio interfiere con una libertad particular:

específicamente con la libertad de percibir la renta bruta. Pero esa no es parte de la libertad de un ser solitario; el superávit representado por la renta sólo existe gracias a la interacción. Y por lo tanto, no es una parte necesaria de la libertad de mercado concebida como una extensión de la libertad natural de la que goza un Robinson Crusoe.

En una sociedad regida por el principio de concesión relativa minimáxima, Chamberlain tiene que renunciar a la alternativa más ventajosa si quiere jugar al básquetbol. Es libre de exigir una remuneración suficiente para hacer que el básquetbol sea su actividad más preferida y libre de negarse a jugar si no ha de obtener esa remuneración. Pero no es libre de perseguir su máxima remuneración posible si ésta incluyera la renta bruta. Sólo tiene derecho a percibir lo suficiente, de modo tal que, además de cubrir el pago de todos sus costos, comparta equitativamente el superávit social del cual la renta bruta es un componente. La pretensión de libertad pesa de una manera en el contrato propuesto por Chamberlain si el jugador exige el pago por los costos de oportunidad y de otra manera muy diferente si el jugador busca obtener un rédito por su talento especial.

Si Nozick considera que el derecho a la renta bruta es un componente de la libertad, John Rawls no sólo exige que se confisque esa renta sino que también pide que se redistribuya de modo tal que, en efecto, el superávit representado por la renta sea disfrutado por aquellos que carecen de dotes ventajosas. Ya citamos antes a Rawls cuando dice: "Los naturalmente beneficiados... deben ganar... sólo para cubrir los costos de la formación y la educación y para emplear sus dotes de modos que ayuden también a los menos afortunados."[10] En otras palabras, cada individuo tiene derecho al pago por los costos de los servicios intermediarios que provee. Pero la renta bruta debe ajustarse al principio de diferencia y por lo tanto debe redistribuirse de modo tal que maximice el bienestar de aquellos que están en peor situación o —para expresarlo de un modo que se aproxime más a la particular formulación que hace Rawls de las exigencias de la justicia— que maximice el nivel mínimo de los bienes primarios sociales y económicos.

Al leer la declaración de Rawls de que "las personas tienen

derecho a ganar para utilizar sus dotes de modos que ayuden a los menos afortunados" parecería que el autor supone que cada persona tiene el derecho de usar su capacidad y su talento como lo desee, por lo tanto que tiene derecho al pago por los costos que surjan de las oportunidades perdidas. Si las personas carecieran de ese derecho, si, por ejemplo, se diera por sentado que cada individuo tiene la obligación de utilizar su capacidad de cualquier modo que maximice el bienestar de aquellos peor situados, nadie tendría pues derecho a un pago por utilizar de ese modo su capacidad y su talento. Cualquier empleo alternativo de las capacidades personales debería quedar descartado por ser contrario a las exigencias de la justicia y por lo tanto no habría costos de oportunidad legítimos que cubrir.

El principio de diferencia léxica que expresa la concepción general de la justicia de Rawls y que establece que "Todos los valores sociales...deben distribuirse equitativamente a menos que una distribución desigual...sea ventajosa para todos",[11] y la idea de que "la distribución de los talentos naturales debe considerarse como un haber común",[12] parecen indicar que ninguna persona tiene derecho a su dote básica de capacidades físicas y mentales que a su vez le dé derecho al pago por los costos de oportunidad. En realidad, ese derecho parece incompatible con la idea de maximizar el bienestar de aquellos peor situados. Pero la especial concepción que tiene Rawls de la justicia, que para él es esa particularización de la concepción general aplicable a nuestras circunstancias sociales y económicas, se expresa en dos principios: el primero y principal exige que "Cada persona tenga un derecho igual al sistema total más amplio de libertades básicas iguales compatible con un sistema similar de libertad para todos".[13] De este principio parecería surgir el derecho a la propia dote básica y por lo tanto el derecho al pago por los costos de oportunidad a cambio de la provisión de los servicios intermediarios .

No seguiremos analizando aquí si la especial concepción de la justicia que sostiene Rawls, que le da a cada persona un derecho a sus dotes básicas, es compatible con su concepción general. Para nuestros fines, lo significativo es que Rawls implícitamente distingue la renta bruta del pago por los servicios intermediarios y coincide con nuestra idea de que cada

persona tiene derecho a este último, pero supone que el primero, en lugar de dividirse equitativamente entre todos los cooperadores, debe utilizarse en beneficio de los menos favorecidos. Y esto constituye una negación de la libertad básica en el contexto de la cooperación. Si cada individuo contribuye igualmente a la producción del superávit cooperativo, cada uno tiene pues derecho a una porción igual de ese excedente. Y Rawls no reconoce este derecho.

4.1 En el capítulo 7 sostuvimos que un individuo podría apropiarse legítimamente de la tierra o de otros bienes que anteriormente fuesen de uso común para su propio uso exclusivo, siempre que el efecto de ese acto fuera que las demás personas quedaran por lo menos tan bien situadas como lo estaban antes. El que se apodera del bien mejora su propia situación sin empeorar las de los demás. Sin embargo, podría objetarse que esa persona cercena las oportunidades de los demás y que por lo tanto empeora la situación de los otros. Al hacer para sí misma la adquisición más ventajosa, la persona que se apropia de un bien que antes era común impide que sus semejantes hagan una adquisición igualmente ventajosa. Alguien podría decirnos pues que debemos tomar más seriamente las palabras del propio Locke "suficiente y tan buena" e insistir en que cualquier oportunidad que un individuo pueda aprovechar legítimamente debe dejar, si no oportunidades similares, oportunidades igualmente ventajosas para los demás. Aunque en realidad una persona pueda quedar en mejor situación que sus semejantes como resultado de apropiaciones no cooperativas, estaría violando la estipulación si los demás no pueden lograr lo mismo que ella.

Pero la función de la estipulación es impedir que alguno saque ventaja de otro. Sin embargo, al aprovechar la mejor oportunidad que se le presenta, un individuo no saca ventaja de los demás, simplemente porque éstos ahora carecen de la oportunidad que aquél aprovechó primero. El ganador de una carrera no saca ventaja de los demás corredores, por lo menos si la participación en la carrera es en sí misma ventajosa. Todos los corredores prefieren una situación en la cual el beneficio sea desigual, de modo tal que haya ganadores y perdedores, a una situación en la cual todos se beneficien equitativamente y no haya carrera.

Si el mundo fuera como se dice que lo conciben los dobu, es decir que existiera una oferta fija de todos los bienes, la estipulación limitaría en realidad la apropiación pues exigiría que se les deje a los demás una oportunidad "tan buena", de modo tal que cada individuo sólo podría apropiarse de una porción igual de la tierra o de los otros bienes. Porque en un mundo semejante las carreras no serían ventajosas para todos los participantes. Por supuesto, en semejante mundo la interacción misma no sería ventajosa, por lo tanto nadie tendría razón para acatar la estipulación. Pero en nuestro mundo, en el cual el esfuerzo humano aumenta la oferta total, particularmente de los bienes materiales, una restricción sobre la apropiación igual sería desventajosa para todos. Aquellos que aprovechan las oportunidades personalmente beneficiosas, con frecuencia también crean nuevas oportunidades como resultado de sus actividades. Por supuesto, el primero de ellos puede obtener el mayor beneficio de su propia actividad. Pero todos pueden razonablemente esperar beneficiarse más de las oportunidades que cree esa actividad, que si el primero hubiese restringido la apropiación de ese bien para dejar a los demás las mismas oportunidades que tuvo él. Ese pionero les deja más a los otros, o mejor dicho, crea más para los otros, si nos les deja literalmente algo "tan bueno". El reconocimiento de esta verdad práctica fundamental ha sido una de las ideas centrales de la sociedad occidental moderna que le permitió alcanzar un éxito nunca antes igualado en la historia humana al dejar que algunos individuos aprovecharan las oportunidades personalmente ventajosas y crearan oportunidades más amplias para los demás.

Pero la estipulación no autoriza toda actividad que cercene ciertas oportunidades a los demás; porque no siempre una de esas actividades ha de crear nuevas y mejores perspectivas. Evidentemente no basta con dejar a los demás tal cual como estaban antes, permitirles continuar con sus actividades presentes y cosechar beneficios iguales a los que obtenían en el pasado, si al mismo tiempo se los priva de las oportunidades con que antes contaban para mejorar sus situaciones sin brindarles nuevas alternativas. Y este punto tiene implicaciones muy significativas si pretendemos determinar el alcance de la apropiación legítima, tanto entre personas como entre grupos.

Consideremos, como lo hicimos en 7.4.1, una isla habitada por personas que originalmente cultivaban la tierra de manera no cooperativa y no pretendían poseer cotos privados. Eva se apropia de una parte de la tierra para su uso personal y cuida de dejarles a los demás lo suficiente para que éstos estén en tan buena situación como estaban antes de que ella se apropiara de esa porción de tierra. Pero luego se descubre petróleo en su terreno. El derecho que ella tiene a esa tierra no necesariamente se extiende al petróleo. Puede ocurrir que todo el beneficio derive de los derechos individuales relacionados con la producción agrícola, que autorizaron la apropiación original de Eva, y que no se extienden a los derechos individuales relacionados con el desarrollo de los recursos minerales o de otro tipo. En este caso, cualquier adquisición individual puede empeorar la situación de alguna otra persona. O puede ocurrir que, aun cuando los derechos individuales a los recursos minerales sean en principio mutuamente beneficiosos, la asignación particular de derechos que resultaría de extender los derechos justificablemente adquiridos a la tierra a los derechos a los recursos de esa tierra, puede no beneficiar a todos.

Eva no puede quejarse justificablemente de que su situación empeore si se le rechaza su supuesto derecho al petróleo hallado en su tierra. Si una persona adquiere una porción de tierra para cultivarla, sin que le importe o sin saber que allí hay petróleo, su posición no puede empeorar porque se le niegue un derecho exclusivo a aquello que se descubra sólo más adelante. Los derechos adquiridos en un acto de apropiación están determinados por el contexto de dicho acto y por el empleo o los empleos que se hagan del bien adquirido. Aquí podemos advertir la verdadera importancia real que tiene lo que dice Locke en cuanto a que lo que se adquiere legítimamente sea utilizado, aun cuando el sentido que el propio Locke le dio a esta exigencia sea defectivo ya que sólo se refería a prohibir el despilfarro.

El punto clave en el que debemos concentrarnos es que el efecto de los derechos al petróleo puede ser muy diferente del efecto de los derechos a la tierra. Al adquirir una porción de tierra, aunque sea la mejor parte de la isla, Eva puede abrir la posibilidad de desarrollar actividades más diversificadas en la comunidad toda, y actividades más especializadas para algunos

individuos particulares, con lo cual se generarían beneficios crecientes para todos. Pero si ella además obtiene derecho al petróleo hallado en su terreno, puede convertirse en la única propietaria del único pozo de petróleo de la isla. Podría llegar a ser la propietaria de un monopolio y la mayoría de las personas quedarían en peor situación si un monopolio privado reemplaza al control común. Por supuesto, si Eva se convirtiera en la propietaria del monopolio no necesariamente tendría que dejar a los demás en peor situación de la que estaban antes de que se descubriera el petróleo, cuando la única actividad era la agricultura. Pero podría estar cercenando en lugar de ampliando las oportunidades de mejorar de los demás al ejercer el control de aquello que podría convertirse en la fuente primaria de energía de todos los habitantes de la isla.

Como los derechos se adquieren definitivamente, al determinar si satisfacen o no la estipulación, es necesario tener en cuenta las consideraciones referentes a sus implicaciones a largo plazo, especialmente en cuanto a la cantidad de oportunidades. La apropiación individualmente beneficiosa no viola en sí misma la estipulación sencillamente porque no satisfaga la literal expresión lockeana "tan buena". Pero si las adquisiciones literalmente desiguales han de formar parte de la posición inicial que cada individuo ha de tener en la interacción de mercado y cooperativa, la apropiación debe darse en un contexto de oportunidades ampliadas.

4.2 El ejemplo esbozado en la subsección anterior puede evidentemente aplicarse al mundo contemporáneo. Los árabes y los albertanos (para mencionar sólo dos casos) afirmaron sus derechos al petróleo que hay bajo sus tierras. Y, según las reglas del actual orden internacional, los reclamos de los árabes, y no los de los albertanos, tienen una base firme: los primeros tienen la buena fortuna de constituir naciones, mientras que los últimos sólo forman una simple provincia. Pero el orden internacional actual no puede brindar una base racional y justa para la interacción cooperativa entre los pueblos del mundo. Los árabes y los albertanos pueden estar ampliando sus propias oportunidades y cercenando las perspectivas de desarrollo de otros pueblos. Por supuesto, ambos pueden sostener ciertas pretensiones históricas que compensen injusticias del pasado.

Pero ésta es otra cuestión y no un fundamento posible para el simple derecho a aquello que pueda encontrarse en el territorio que casualmente uno ocupa.

Los derechos asignados a las personas del orden internacional —a las naciones— abren la puerta a perspectivas de oligopolios y monopolios que ejerzan el control, lo cual bien puede constituir una violación de la estipulación lockeana, cuando se tiene en cuenta el efecto que producen en las oportunidades para el desarrollo. Pero aquí podemos advertir algo que parece constituir una violación más seria, y que surge no de reconocer los derechos que legitiman el acceso directo de los pueblos a los recursos de las tierras que ocupan, sino del uso de los recursos sobre los cuales otros pueblos reconocen su derecho, aun cuando ese uso no siempre se haga mediante una coerción directa o abierta.

Si una nación se industrializa y goza de la opulencia, mientras otras no se industrializan y permanecen en el nivel de la agricultura de mera subsistencia, no necesariamente existe injusticia. Un país industrializado puede respetar plenamente la estipulación lockeana, enriqueciéndose de modos que creen nuevas perspectivas para sus vecinos, sacando o no ventaja de estos últimos. Pero si esa nación industrializada intenta mediante la coerción o el acuerdo, apropiarse de los recursos por el momento no utilizados de sus vecinos subdesarrollados, de modo tal que éstos ya no estén en posición de elegir sus propias vías de desarrollo sino que antes bien deben decidir entre continuar con su estilo original de vida o proveer a las necesidades de su socio industrializado —produciendo para él café, bananas, carne, caucho o celulosa—, en ese caso, pues aun cuando el efecto inmediato de la asociación entre ambos sea aumentar el estándar de vida y ampliar las oportunidades de progreso individual en las naciones subdesarrolladas, el efecto a largo plazo muy probablemente disminuya tales oportunidades pues condenaría a los países subdesarrollados a la dependencia y el empobrecimiento permanentes. Si una nación puede sacar ventaja de otras adquiriendo un monopolio efectivo sobre ciertos recursos de gran necesidad, también puede ser víctima de otras que exploten los recursos que a largo plazo constituyen la clave de su propio desarrollo.

Evidentemente los vecinos subdesarrollados de economía agrícola no pueden hallarse en posición de desafiar a su socio industrializado; pues este último sencillamente dicta los términos de la interacción. En las transacciones a corto plazo por lo menos, es probable que se dé la depredación. Pero con esto no estamos diciendo que queden legitimados los derechos *de facto* que constituyen el punto de base de esa interacción. El acceso diferenciado a los recursos y los diferentes niveles de desarrollo son factores que pesan contra la tendencia a acatar la estipulación en el actual estado de las naciones. Señalar esto y mostrar que la base de cooperación internacional existente es en parte injusta, puesto que aquellos menos favorecidos por tal cooperación sólo pueden aceptarla racionalmente en la medida en que se la mantenga coercitivamente, no significa acusar de perversidad deliberada a aquellos que se benefician por tener mayor acceso al desarrollo o por gozar de un desarrollo mayor. En realidad, mientras el estado de las naciones continúe siendo un estado de naturaleza, aquellos que medran a expensas de sus vecinos se conducen como maximizadores directos en circunstancias en las cuales la maximización restringida no recibiría una respuesta recíproca. La estipulación sólo es una restricción racional de la interacción entre las naciones si existe la perspectiva de un orden internacional de cooperación. Pero debemos recordar que la distribución natural *de facto* que resulta de la interacción de las naciones, y que es un estado muy semejante al de la naturaleza no puede determinar una asignación aceptable de derechos como una base para relaciones más cooperativas. Si comparamos esos derechos con la estipulación, debemos considerar el aumento y la ampliación mutua de oportunidades como un criterio fundamental para juzgar qué puede pretender razonablemente cada socio como dote inicial en un orden mundial justo.

4.3 El encuentro entre dos pueblos que se hallan en diferentes niveles de desarrollo da lugar a cuestiones referentes a la base y el carácter legítimos de su interacción que se extienden más allá de la cuestión particular de las oportunidades. Imaginemos un planeta cuya superficie estuviera cubierta de mares tormentosos sólo quebrados por dos grandes islas ampliamente separadas entre sí. Cada isla está habitada por

criaturas de aspecto suficientemente parecido al de los humanos que podríamos considerarlas de nuestra propia especie; una isla está habitada por la gente púrpura y la otra por la gente verde. Antes de que comenzara nuestra historia, los dos pueblos habían vivido en completo aislamiento mutuo. Ni la gente púrpura ni la gente verde tenía conciencia de que existieran otros seres humanos en su planeta.

La gente púrpura había desarrollado una sociedad idealmente justa. Los derechos de cada persona púrpura estaban sólidamente afirmados en la estipulación. Un mercado libre extendía su dominio y donde lo hacía conseguía la optimidad; además las externalidades estaban contrarrestadas por medidas políticas y prácticas que satisfacían el principio de concesión relativa minimáxima. A fin de garantizar la equidad a largo plazo (un problema que examinaremos más detalladamente en la última sección de este capítulo), se habían tomado medidas políticas referentes a la población, a la conservación y al desarrollo destinadas a proveer a cada persona de beneficios equivalentes —en cuanto a sus derechos, a sus características personales y a las formas posibles de interacción social— a aquellos que recibiera cada una de las demás personas. La gente verde, por su lado, vivía en la miseria y en un completo caos. Incapaces de interesarse por sus semejantes y por su propio futuro, los verdes malgastaban sus recursos, avanzaban sobre los demás sin restricciones, de modo tal que se hallaban en el límite de un catastrófico colapso. En ese momento de sus respectivas historias, un grupo de exploradores del país de los púrpuras descubre a los verdes y de regreso informa sobre las condiciones en que viven estos últimos a sus compañeros de isla.

La gente púrpura se pone de acuerdo en varias cuestiones. En primer lugar, cualquier contacto que se realice entre los dos grupos deberá hacerse por obra de la instigación de los púrpuras, ya que los verdes carecen tanto de medios de transporte como de comunicación a través del mar. Los recursos combinados de las dos islas no pueden mantener a las poblaciones combinadas en el nivel de bienestar de que disfrutan los miembros de la sociedad púrpura. Maximizar el promedio o el nivel mínimo de bienestar exigiría una transferencia masiva de recursos de la gente púrpura a la gente verde.

La gente púrpura no tiene razones para adherirse a una respuesta utilitarista ante esta situación que acabamos de esbozar, una respuesta que obligara a los púrpuras a emprender una transferencia masiva de recursos a los verdes. Por supuesto, los púrpuras pueden responder por simpatía a los aprietos que sufren los verdes. Hasta pueden considerarse obligados a no empeorar la desafortunada situación en que viven los verdes, aunque si nuestra teoría puede sostener tal obligación es una cuestión que debemos considerar. Pero los púrpuras no tienen, ni pueden tener, razones para considerarse obligados a empeorar su propia situación con el fin de maximizar el promedio general de utilidad.

Es evidente que los púrpuras no violaron la estipulación respecto de los verdes. No pudieron hacerlo puesto que no existió ninguna interrelación en el pasado. Por consiguiente, en la situación actual de los dos pueblos no existe una injusticia, a pesar de que entre las condiciones de ambos haya desigualdades de mercado. La injusticia sólo puede darse en la medida en que las acciones de una persona o de un grupo afecten a otra u otras. Por lo tanto, la situación existente constituye una posición inicial aceptable para aplicar los principios de negociación racional. Las desigualdades que forman parte de esta situación han de conservarse en cualquier acuerdo cooperativo racional que se realice entre los púrpuras y los verdes. Tales desigualdades no forman parte del superávit cooperativo que haya de distribuirse luego entre los dos pueblos de acuerdo con los términos de la cooperación.

¿Cómo podemos esperar que se divida el superávit que surja de la cooperación? Teniendo en cuenta los niveles de desarrollo tan diferentes alcanzados por los dos pueblos y las posiciones muy diferentes que ambos ocupan, la cooperación de púrpuras y verdes con toda probabilidad no ha de proporcionar un bien único y transferible que pueda repartirse igualmente entre ellos. En la medida en que los beneficios de la cooperación sean comparables, parecería evidente que los verdes tienen mucho más que ganar de la interacción entre ambos que los púrpuras. De ahí que si el superávit cooperativo se divide en porciones equivalentes, como lo exige la concesión relativa minimáxima, los verdes han de obtener beneficios absolutos mayores.

Por lo tanto podría suponerse que, partiendo de una posición de gran desigualdad, los dos pueblos gradualmente irían acercando sus niveles de desarrollo y bienestar. Pero no necesariamente ése debe ser el caso; la desigualdad absoluta entre los dos pueblos hasta puede aumentar, aunque la cooperación beneficie más a los verdes que a los púrpuras. Porque, como lo mostraremos en la última sección de este capítulo, la cooperación racional con una sociedad debe dar en general por resultado una tasa positiva de crecimiento y, por lo tanto, un aumento del bienestar material a lo largo del tiempo. Posiblemente esa tasa interna de crecimiento sea tanto mayor entre los púrpuras que entre los verdes que sus efectos compensen con exceso los mayores beneficios que los verdes reciben de la cooperación. Un aumento de las desigualdades existentes entre dos pueblos que cooperan no es una señal de que los términos de la cooperación entre ambos sean injustos ni de que contraríen las exigencias de la concesión relativa minimáxima.

Nuestra teoría moral parece directamente aplicable a la interacción que se desarrolle entre púrpuras y verdes. La cooperación comienza partiendo de la situación existente cuando los dos pueblos se pusieron en contacto por primera vez; la cooperación misma tiende a beneficiar más al grupo menos desarrollado, porque éste tiene más que ganar; a pesar de ello, la cooperación a lo largo del tiempo no tiene por qué llevar a ninguna reducción de las desigualdades que existían ya entre los dos grupos. Pero pueden darse algunas complicaciones que no hemos considerado aún. ¿Los púrpuras tienen necesidad de entablar la cooperación con los verdes?

Algunos de los púrpuras podrían muy bien oponerse a cualquier forma de interacción con sus congéneres recién descubiertos. Podrían hacerlo aun cuando esperaran beneficiarse, aunque sólo sea marginalmente, de los intercambios con los verdes. Porque supondrían que esas ganacias estarían acompañadas por el grave costo de mantener una relación que, aunque justa, con el tiempo podría despertar el resentimiento de los verdes en virtud de su desigualdad. Sin duda las personas idealmente racionales no manifestarían semejante resentimiento, pero los verdes no son idealmente racionales. Además, los púrpuras podrían suponer que una relación en alto grado

desigual forzaría sus propios sentimientos, es decir que la simpatía los impulsaría no sólo a compartir los beneficios de la cooperación sino también a transferir bienes propios a los verdes menos afortunados. Y los púrpuras podrían preferir no situarse en una posición en la cual sus propios sentimientos de simpatía, combinados con el resentimiento de los verdes, pudieran colocarlos en una situación peor de la que disfrutarían si no se diese ninguna interacción y por lo tanto no surgiesen ni sentimientos de simpatía ni resentimientos. En efecto, puede sostenerse que en esta situación en alto grado desigual, los sentimientos hacia los demás difícilmente constituyan una fuente de enriquecimiento de la vida humana.

Si algunos de los púrpuras prefieren ignorar a los verdes, otros pueden sugerir que los púrpuras hagan el uso que puedan de la isla habitada por los verdes, sin causarles a éstos daño o desventaja deliberada y directa y sin preocuparse tampoco por los efectos indirectos que pueda tener ese uso. O, de manera más radical, pueden proponer que los púrpuras se lancen activamente a la depredación y tomen lo que puedan obtener de los verdes y de la isla habitada por éstos. La segunda de estas proposiciones viola intencionalmente la estipulación. La primera, en cambio no lo hace, en la medida en que los púrpuras no pretenden obtener beneficios de la interacción con los verdes, puesto que tal interacción sólo les provoca la expectativa de utilidad que ellos obtendrían si la otra isla estuviera deshabitada. La preferencia por una de estas dos proposiciones dependerá de las circunstancias particulares; también podría ocurrir que los púrpuras decidieran adoptar una política directamente maximizadora de la utilidad. En efecto, sus defensores afirman que los púrpuras no tienen ninguna buena razón para extender su disposición a la maximización restringida, de modo tal que abarque también la interacción con los verdes.

Podríamos caracterizar esa política no cooperativa diciendo que es una forma de tratar a los verdes no como seres humanos semejantes, no como "nosotros", sino como animales, como "ellos". La conducta con los animales es por completo maximizadora de la utilidad, aunque los sentimientos particulares por ciertos animales puede afectarla. Al basar la moral en la elección racional, excluimos de la esfera de la restricción

moral aquellas relaciones con criaturas no humanas. ¿Por qué los púrpuras no podrían considerar a los verdes como meros animales?

Hay varios terrenos en los cuales los púrpuras podrían defender la maximización restringida, la adhesión a la estipulación lockeana y la aceptación de la concesión relativa minimáxima en sus relaciones con los verdes. En primer lugar, podrían afirmar que, sean cuales fueren las ventajas a corto plazo, a la larga los púrpuras pueden esperar beneficiarse con la restricción. A medida que continúe la interacción entre los púrpuras y los verdes, estos últimos pueden llegar a mejorar la forma de defender sus propios intereses. Entonces la cooperación, en términos justos, puede llegar a ser ventajosa para ambas partes. Pero si los púrpuras se conducen como depredadores o sencillamente hacen caso omiso de los verdes, es probable que deban afrontar la implacable hostilidad de estos últimos, o en el mejor de los casos una demanda por una compensación, como una condición previa a la cooperación y esa compensación puede implicar costos mayores de los que hubiera exigido una adhesión continuada a la estipulación. Reinstaurar una posición inicial justa para la cooperación puede ser menos beneficioso que no apartarse nunca de ella. Por lo tanto, puede sostenerse que la ventaja esperada exige a los púrpuras disponerse a cooperar con los verdes.

Pero puede darse una segunda posibilidad en la que los púrpuras ya estuvieran dispuestos a tal cooperación. Al desarrollar su propia sociedad idealmente justa, éstos pueden haber introducido pautas de socialización que los dispusieran a la maximización restringida con todos aquellos a quienes consideren seres humanos semejantes y puede que no tuvieran otra alternativa que considerar de ese modo a los verdes. Digamos que están dispuestos a tratar a todos los seres humanos con ciertas aptitudes y actitudes como socios potenciales de los arreglos cooperativos y que los verdes puedan tener esas aptitudes y actitudes. Quizá si hubiesen sabido que alguna vez se encontrarían con los verdes, los púrpuras se habrían dispuesto a tratarlos de manera diferente. Pero aunque así fuera, no necesariamente tienen razones para intentar cambiar sus pautas de socialización. Los beneficios y costos que reporta establecer ciertas pautas pueden ser muy diferentes de aquellos que exigen alterar tales pautas. Por lo tanto, los púrpuras, sin demostrar por ello ninguna irracionalidad, pueden simplemente considerar la posibilidad de no acercarse nunca a los verdes y permanecer en una posición no cooperativa. Por supuesto, los púrpuras no habrían de comprometerse en una cooperación activa salvo en el contexto de un acuerdo explícito o implícito, pero pueden mostrar desde el principio su disposición a tal acuerdo adhiriéndose estrictamente a la estipulación.

Y en tercer lugar, puede darse que los púrpuras sientan un cierto grado de simpatía por todo aquel a quien consideren un ser humano. Pueden sentir un interés positivo por los intereses de los demás y esto puede bastar para fundamentar una disposición a no empeorar la situación de otras personas, siempre que la alternativa no sea aceptar un empeoramiento significativo de la propia situación. En ese caso posiblemente los púrpuras no deseen apropiarse de los recursos de la isla recién descubierta sin preocuparse por los aprietos que sufren los verdes. Sin reconocer ninguna obligación positiva hacia los verdes y sin aceptar ningún deber que les exija renunciar a alguno de los beneficios de los que actualmente gozan a fin de mejorar la desafortunada situación de los verdes, los púrpuras pueden considerarse obligados, en virtud de sus sentimientos de simpatía, a no hacer nada que empeore la situación de sus vecinos. Aunque los púrpuras no tengan la menor razón para reprocharse por los actuales infortunios de los verdes, podrían reprocharse si hicieran algo que afectara a éstos adversamente.

Una reacción semejante no es únicamente una exigencia de la razón, ni puede serlo. Las demandas de simpatía son algo por completo diferente de aquellas de la elección racional y tratarlas como algo equivalente sólo puede llevarnos a confusión. Una moral de los sentimientos debe analizarse independientemente por completo de una moral basada en el acuerdo. Pero podemos suponer que para ciertas personas el hecho de disponerse a la interacción cooperativa antes que a conducirse como un depredador y por lo tanto a restringir la maximización de la propia utilidad de acuerdo con la estipulación y el principio de concesión relativa minimáxima, puede basarse en sentimientos de simpatía.

Al despedirnos de la gente púrpura y la gente verde, podemos relacionar esa ficción con la realidad. Las relaciones que mantienen los países más desarrollados con los menos desarrollados en nuestro mundo real no tienen en muchos aspectos significativos un paralelo con las relaciones que podrían mantener los púrpuras y los verdes. Pero por lo menos algunos estudiosos de la moral contemporáneos sostienen que, aunque tuvieran un paralelo, deberíamos considerar injusta la situación y defender la idea de que los países desarrollados tienen la obligación de hacer sacrificios para beneficiar a los menos desarrollados. Porque a veces se supone que el mero hecho de que exista la desigualdad —y el mero hecho de que algunas personas gocen de un bienestar mucho mayor que el de otras y que esa posición mejore mucho más velozmente— es una señal de injusticia. Nuestro relato debería poner de manifiesto la falacia de semejante opinión. Podría darse el caso de que Laos o Mali, Burkina, Fasso o Etiopía fueran aun más pobres y que los Estados Unidos o Suiza fueran aun más ricos si todos estos países fueran islas diferentes que no mantuvieran contacto entre sí. Podría ocurrir que las desigualdades aumentaran, no porque los que saben hacer mejor las cosas se aprovechen de los más pobres, sino porque la tasa interna de crecimiento que las naciones ricas pueden mantener justamente excede la que es posible para las pobres y supera los beneficios que procura la cooperación a los países menos desarrollados.

Con esto no estamos diciendo que el orden internacional actual sea un modelo de cooperación justa y racional. Estamos bastante dispuestos a suponer que por lo menos algunas de las desigualdades y disparidades existentes derivan de la explotación, de una interacción que viola la estipulación y la concesión relativa minimáxima. Pero no podemos suponer esto sencillamente porque tales disparidades existen. Debemos preguntarnos cómo llegaron a darse. Si pretendemos que las naciones mantengan entre sí una interacción racional, justa y no explotadora, es necesario determinar si las relaciones que mantuvieron en el pasado no incluían violaciones a la estipulación que obliguen a una compensación y a ajustar los derechos *de facto* existentes. También es necesario determinar si las formas actuales de interacción benefician equitativamente a cada una

de las naciones. Hasta que podamos responder a estas preguntas, por supuesto no con precisión sino con una aproximación razonablemente segura, no estaremos en posición de evaluar el orden internacional actual, por más que ciertos hechos, o supuestos hechos, puedan conmover nuestros sentimientos. En el mundo de los púrpuras y los verdes las desigualdades están justificadas; en el nuestro quizá lo estén o quizá no. Aún no lo sabemos.

5.1 La interacción entre pueblos que mantienen diferentes estilos de vida da lugar a ciertas cuestiones de gran complejidad práctica que nuestra teoría moral puede ayudar a esclarecer. A lo largo de toda la historia humana, el contacto entre los diferentes estilos de vida demostró una tendencia a provocar conflictos y, con frecuencia, conquistas, antes que una coexistencia pacífica. La idea de que cada estilo de vida tiene derecho a conservarse —una idea expresada en esa extraordinaria concepción que conocemos como genocidio cultural— es una recién llegada al escenario de la moral. Y además es una idea completamente extraviada. Los individuos importan; los estilos de vida importan sólo como expresión y como alimento de la individualidad humana. Basándonos en nuestra teoría y en unos pocos supuestos fácticos pero profundamente controvertidos sobre lo que los seres humanos prefieren, podemos llegar a comprender más claramente las circunstancias en las cuales un estilo de vida reemplaza de manera justificada a otro.

Comenzaremos por formular cuatro criterios para establecer que un estilo de vida es más avanzado que otro, que ha alcanzado un estadio superior de desarrollo humano. El primero, quizás el menos importante, es la *densidad de población*: cuando los demás factores son iguales, el estilo de vida *A* es más avanzado que el *B* si permite a un número mayor de personas habitar un determinado territorio. El segundo criterio es la *duración de la vida*: cuando los demás factores son iguales, el estilo de vida *A* es más avanzado que el estilo de vida *B* si permite que aquellos que lo llevan gocen de un promedio de vida más largo. El tercer criterio es el *bienestar material*: cuando los demás factores son iguales, el estilo de vida *A* es más avanzado que el *B*, si permite que quienes lo llevan gocen de una mayor riqueza promedio y de una variedad mayor de bienes materia-

les. El cuarto y más importante criterio es la *amplitud de las oportunidades*: siendo los demás factores iguales, *A* es un estilo de vida más avanzado que *B* si aquellos que lo llevan disfrutan en promedio de una mayor posibilidad de elección entre funciones profesionales y no profesionales diversas y variadas. Estos cuatro criterios determinan un ordenamiento parcial de los estilos de vida, desde el punto de vista del progreso o el desarrollo humano. Y no intentaremos asignarles a estos criterios más fuerza que la de proporcionar un ordenamiento parcial.

¿Cómo podemos defender tales criterios? Apelamos a lo que parecen ser hechos evidentes relativos a las preferencias humanas. Aunque no todos los seres humanos deseen reproducirse, la mayor parte de ellos lo desean y ciertamente la mayor parte desea tener la oportunidad de hacerlo. Una sociedad que mantiene a una población mayor, con promedios de vida más largos, les brinda a las personas oportunidades más efectivas de poder reproducirse. Y también, aunque no todos los seres humanos deseen prolongar sus vidas, la mayor parte de ellos lo desea y también lo desea para sus parientes y amigos. Una sociedad que ofrece un promedio de vida más largo da una respuesta positiva a ese deseo.

El deseo de alcanzar el bienestar material se manifiesta primariamente entre aquellos que lo creen posible, aquellos que consideran que el mundo es un lugar de riquezas crecientes. Entre aquellos que ven así el mundo, el deseo se amplía y se amplía aun más el deseo de tener la oportunidad de aumentar su bienestar material. Una sociedad que proporciona un nivel promedio superior de bienestar material da una respuesta positiva a ese deseo; les brinda a sus miembros una opción de la que carecen los miembros de una sociedad materialmente pobre.

Además, el deseo de contar con una cantidad mayor de oportunidades profesionales y no profesionales, entre las cuales poder elegir, se manifiesta primariamente entre las personas que tienen una formación social que les hace concebirse como individuos libres y autónomos y no predestinados desde su nacimiento a desempeñar cierto papel social predeterminado. Los miembros de un pueblo primitivo de cazadores y recolectores de los frutos de la tierra no conciben la idea de poder elegir entre

diversas oportunidades. Pero en este caso el desarrollo de la preferencia es unidireccional; a medida que las personas cobran conciencia de que existe un campo mayor de oportunidades, desean ser libres de elegir entre ellas; a medida que las personas cobran conciencia de que el campo de oportunidades se estrecha, lamentan la pérdida de la libertad que conlleva ese estrechamiento. Una sociedad que amplía la cantidad de oportunidades efectivamente accesibles provee cada vez mejor a las variadas aptitudes y destrezas, las variadas preferencias y actitudes que caracterizan a los seres humanos. En 8.4.2 vimos que la elección fundamental del actor ideal es poder utilizar su capacidad para satisfacer lo más posible sus preferencias. Pero entonces el actor ideal debe también elegir una sociedad que les brinde a sus miembros la posibilidad más amplia de elección entre los papeles profesionales y no profesionales que se puedan desempeñar, la gama más amplia de oportunidades y los medios más efectivos para realizar esa elección fundamental.

Los cuatro criterios nos dan una perspectiva normativa en la cual podemos estimar muchas de las interacciones que se dan entre pueblos que llevan diferentes estilos de vida. No habrá escapado a la atención del lector ni —a este punto de nuestra argumentación— lo habrá sorprendido, advertir que entre las sociedades humanas, la que satisface preeminentemente los criterios es nuestra propia civilización occidental moderna. Basta un sólo ejemplo para ilustrar esto. A comienzos del presente siglo, la duración promedio de la vida en el nivel mundial se situaba entre los 30 y los 40 años.[14] Ochenta años después, ese promedio se elevó a 60 años. Puesto que la edad mínima hipotética para los seres humanos es de alrededor de 20 años (la especie no puede reproducirse efectivamente y por lo tanto no puede sobrevivir, si el promedio de duración de la vida es menor), en este siglo el lapso de vida de hombres y mujeres aumentó por lo menos tanto (y quizá significativamente más) como lo que creció en general en toda la historia humana previa. Puede atribuirse primariamente ese aumento a las específicas ideas sanitarias occidentales y a la tecnología que instrumentó tales ideas, complementada por progresos en el campo de la nutrición, de la productividad agrícola y del transporte específicamente occidentales. Sea cual fuere la motivación de

los mercaderes, los soldados y los misioneros que partieron de Europa y de Norteamérica, el contacto que establecieron tales viajeros con los demás pueblos del mundo condujo finalmente a duplicar las expectativas de vida de esos otros pueblos. Como un legado imperial, lo que se dio fue un empeoramiento de la situación.

5.2 Volvamos a hacer una visita final a nuestra isla imaginaria. Supongamos que Eva, al apropiarse de una parte de la tierra para su uso exclusivo, toma una porción tan grande que les impide a los demás vivir tan bien como lo hacían antes si continúan utilizando la tierra que les queda como pastos comunes. Sin embargo, Eva les proporciona oportunidades efectivas de aumentar su bienestar si ellos deciden cambiar su estilo de vida. En el terreno que adquiere, Eva introduce nuevas técnicas de producción alimentaria y así aumenta el rendimiento en tal medida que algunos de sus vecinos pueden dejar de ser autosuficientes y especializarse en cambio en oficios artesanales que les permitan intercambiar sus productos por la comida excedente de Eva y en consecuencia vivir mejor. Los vecinos que prefieren continuar cultivando la tierra pueden aprender a dominar las nuevas técnicas que Eva comparte libremente con ellos, dividir la tierra común en predios individuales y poner en práctica esas nuevas técnicas con el fin de beneficiarse tanto por las mayores cosechas como por el intercambio con los artesanos. En consecuencia, aunque los demás habitantes de la isla ya no gocen de la libertad de continuar con sus viejos métodos sin sufrir pérdidas, todos pueden beneficiarse cambiando esos métodos y recibir, en virtud de las oportunidades ampliadas que les ofrecen nuevos y mayores beneficios, una compensación por la pérdida de los beneficios que les aportaban sus actividades originales.

Esa compensación, ¿basta para satisfacer la estipulación? Eva mejora su situación en relación con, y en realidad mediante la interacción con, sus vecinos. Al hacerlo les brinda las oportunidades de mejorar sus situaciones, pero sólo si ellos aprovechan tales oportunidades pueden evitar encontrarse en peor situación de la que estaban antes de que Eva se apropiara de una porción de la tierra. Podemos suponer que al principio, los vecinos de Eva prefieren no cambiar sus estilos de vida ni

aprovechar las nuevas oportunidades. Pues sus preferencias se relacionan con las antiguas circunstancias. Pero también podemos suponer que, de acuerdo con el análisis que hicimos en la subsección anterior, si los vecinos de Eva cambian, no han de querer regresar a sus antiguos estilos de vida, no meramente porque adquirieron nuevos objetivos y nuevos deseos que el antiguo estilo de vida no puede satisfacer, sino porque prefieren el nuevo estilo de vida que les brinda más oportunidades y un bienestar material mayor al estilo que, según ellos suponen, habrían llevado si no hubiesen cambiado. Los demás habitantes de la isla han de encontrarse no en la posición del fumador habitual que no puede prescindir de un cigarrillo pero desea no haber comenzado nunca a fumar, sino en la posición del ex obeso que logró perder peso, que soportó la agonía de someterse a dieta pero que descubre que con su cuerpo delgado disfruta de una vida más feliz.

Eva ofrece dos de los criterios de desarrollo humano que establecimos antes. Les ofrece a sus vecinos la oportunidad de elegir entre diferentes funciones profesionales —agricultor, artesano, constructor— y la aumentada riqueza, producto del intercambio entre esas personas que cumplen las diferentes funciones. Por supuesto, la expansión y la diversificación no pueden darse sin cierta contracción. Las oportunidades de fabricar automóviles o de dedicarse a la mecánica desplazan a las oportunidades de los cuidadores de caballos y los herreros, para mencionar un ejemplo clásico. Semejante contracción representa el costo inevitable del desarrollo humano. Sin embargo, si no se excluyera a los individuos de la expansión general de las oportunidades (con el correspondiente aumento del bienestar material) ni se los obligara a formar parte de ella, no se empeoraría la situación de ciertas personas a fin de mejorar la de otras y en justicia nadie podría quejarse del cambio. Con lo cual queda reivindicada la apropiación de Eva.

La estipulación no exige que el individuo deje a los demás tan bien situados como estaban antes realizando sus anteriores actividades, si les suministra nuevas y efectivas oportunidades de mejorar esa situación. Hasta aquí hemos supuesto que esas oportunidades son optativas; aquellos a quienes se les presentan pueden elegir entre cambiar o continuar siendo como eran.

¿Esa elección es en sí misma necesaria para evitar que se viole la estipulación? ¿O podemos sostener que mientras los demás puedan beneficiarse mediante un cambio de sus estilos de vida, no es injusto obligarlos al cambio? ¿ Puede Eva forzar a algunos de sus vecinos a convertirse en artesanos y a otros en terratenientes independientes, en lugar de permitirles decidir si prefieren continuar ganándose la vida con los pastos comunes?

Nuestra respuesta a esta última pregunta debe ser firmemente negativa. Porque el hecho de que Eva obligue a los demás a cambiar constituiría una violación al derecho que cada persona tiene de usar su cuerpo y su capacidad como mejor le convenga, siempre que al utilizarlos no viole la estipulación. Eva puede brindar oportunidades de cambio y en realidad debe brindarlas para justificar su apropiación, pero no puede forzar a los demás a aceptar tales oportunidades sin violar los derechos básicos de sus vecinos. Puesto que el derecho a la propia persona constituye el primer paso para transformar el estado de naturaleza hobbesiano, no puede revocárselo luego. Por lo tanto, la apropiación sería injusta si no les permitiera a los demás vivir como antes, aunque eso signifique vivir peor. El derecho a adquirir una porción de tierra u otros bienes para darles un uso exclusivo no implica el derecho a obligar a los demás a participar de una interacción contra su voluntad, aun cuando esa interacción les sea beneficiosa.

Antes de seguir adelante deberíamos observar que las oportunidades que se les brinden a los demás de mejorar sus situaciones deben ser efectivamente accesibles. Eva debe proporcionarles a sus vecinos toda la ayuda necesaria para que éstos conozcan y puedan aprovechar tales oportunidades, pues ellos deben elegir entre adoptarlas o sufrir una pérdida. La importancia práctica que tiene este punto para poder aplicar la estipulación resulta evidente.

Ahora cambiemos nuestro relato en un aspecto significativo. Supongamos que un extranjero, a quien llamaremos Colón, aparece en nuestra isla cuando ésta constituye un bien común de todos sus habitantes. ¿Puede Colón unirse a los habitantes originales y explotar la tierra con ellos? Podemos suponer que la llegada del intruso empeora la situación de los nativos, pues crea una presión poblacional mayor sobre la tierra. Sin embargo, si Colón es enteramente autosuficiente y no se aprovecha de ningún modo de la presencia de los pobladores originales, no mejora su propia situación en comparación con lo que obtendría en ausencia de los nativos. No viola la estipulación al hacer uso de la isla, aunque igualmente, los pobladores originales no violan la estipulación si deciden excluirlo. Puesto que al excluirlo no mejoran su situación en comparación con lo que obtendrían en ausencia de Colón.

No obstante, si Colón trata de entrar en interacción con los nativos y de beneficiarse con la presencia y las actividades de éstos, la situación es por completo diferente. Si Colón meramente utiliza la tierra, a menos que se aproveche del cultivo previo de ésta, no se beneficia de ningún modo con lo que hacen los demás. Pero si en cambio les propone comerciar, el costo que tiene para los pobladores originales la presencia de Colón entra dentro del marco de la interacción como un costo desplazado, como una externalidad. Entonces aparece la estipulación. Salvo en el caso de que Colón se proponga obrar como un depredador, debe aceptar la posición inicial determinada por la estipulación. Y puesto que ésta exige la internalización de todos los costos, Colón debe brindarles a los pobladores originales beneficios suficientes para compensarlos por su presencia. Sólo una vez que haya hecho esto habrá sentado una base justa para el intercambio.

Supongamos que el principal beneficio que Colón aporta es una tecnología agrícola más avanzada. Y supongamos además que lo que busca Colón es no compartir la tierra con los pobladores originales sino apropiarse de una parte para su uso exclusivo, de modo tal de poder instrumentar una revolución agraria. En ese caso el intruso debe empeorar la situación de los habitantes originales si éstos deciden continuar como antes, puesto que quedan confinados a un terreno menor. Pero el mayor rendimiento de la tierra que cultiva le permite a Colón desempeñar el papel que le tocaba a Eva en nuestro relato anterior. El recién llegado produce un superávit de alimentos y los intercambia voluntariamente con sus vecinos por otros bienes que algunos de ellos pueden ahora especializarse en producir, con lo cual se obtienen más oportunidades y un mayor bienestar para todos. Los pobladores originales deben elegir

entre su antiguo estilo de vida, que implica un bienestar disminuido, y el nuevo estilo que les ofrece Colón, que implica horizontes más amplios y mayores beneficios materiales.

Colón es un extranjero que toma algo que siempre fue utilizado exclusivamente por otros. Los habitantes originales pueden alegar que la isla les pertenece, que ellos la utilizan plenamente para mantenerse y que el intento de Colón de apropiarse de una porción de la isla para su uso exclusivo es un robo. A su vez Colón puede alegar que en realidad los pobladores originales tienen un derecho en la isla, plenamente reconocido en las circunstancias de su apropiación, pero no un derecho a la isla. Desde su punto de vista, una parte de los recursos y de la capacidad de la isla se está derrochando. Mucha más gente puede habitar esa isla y cada una de esas personas puede gozar de un nivel material más alto si adopta su tecnología más avanzada. Una vez más debemos apelar a los criterios enumerados en 5.1. Y así como los derechos adquiridos en un acto de apropiación están determinados por el contexto del acto mismo y por el uso que se le dé al objeto de la apropiación, del mismo modo las limitaciones establecidas sobre la apropiación por el uso previo están determinadas por la naturaleza de ese uso en relación con los criterios del progreso. Si un acto de apropiación se da en el contexto de usos nuevos que permiten a una mayor cantidad de gente vivir con un bienestar material mayor, no necesariamente implica una violación a la estipulación, aun cuando la persona que se apropie del bien sea un extranjero.

En el estado de naturaleza, si no ya en cualquier sociedad, el uso eficiente es una condición de la posesión legítima. Aunque los derechos sean permanentes, están sujetos a la alteración que producen las nuevas posibilidades. La importancia que tiene el uso eficiente para que exista una posesión legítima dentro de una estructura social establecida depende por supuesto de los términos implícitos o explícitos del acuerdo en el que se fundamenta esa estructura. De ahí que en un tribunal legal la defensa que hace el extranjero al justificar su apropiación mediante los nuevos usos que posibilitan el progreso, puede no resultar válida. Pero, en el tribunal de la naturaleza, en el cual lo único que el individuo debe mostrar es que no empeora la situación de los demás al tratar de mejorar la suya mediante la interacción

con los otros, el extranjero puede justificar su apropiación si les brinda a aquellos cuyas actividades presentes afecta adversamente, perspectivas nuevas y más amplias que les permitan aumentar su bienestar. De modo que queda así reivindicada la apropiación de Colón, como lo fue la de Eva.

5.3 Al llamar Colón a nuestro extranjero sugerimos la relevancia que tiene nuestra historia para evaluar las relaciones que se dieron entre los pobladores nativos de Norteamérica y los primeros europeos que entraron en contacto con ellos.[15] Sostendremos que gran parte de la apropiación que los europeos hicieron de América del Norte podría haberse realizado sin violar la estipulación y, por lo tanto, sin tratar injustamente a los nativos. No sostendremos que lo que realmente ocurrió estuvo exento de injusticia. Sin embargo, a los fines del presente análisis es importante imaginar una apropiación hipotéticamente justa. Si podemos mostrar cómo, sin cometer injusticias, los europeos podrían haber habitado Norteamérica con derechos a las tierras que ocuparon, derechos válidos aun contra los reclamos hechos por los pobladores nativos, sería posible reparar las injusticias realmente cometidas sin exigir que los europeos renunciaran a sus derechos existentes. Si, por otro lado, no pudo existir para las europeos una manera de ocupar Norteamérica, salvo la de apropiarse de lo que encontraban como depredadores, lo cual no les daba ningún derecho y con lo cual violaban la estipulación (si era eso lo que ofrecían para determinar la posición inicial de una interacción cooperativa con los habitantes nativos) no sería posible ninguna reparación efectiva y la cooperación justa exigiría una base muy diferente de la verdadera situación presente.

Imaginemos pues el continente habitado por indios y por inuit que lo utilizaban como vastos pastos comunes. Supongamos que el nivel alcanzado por la población hubiese sido tan alto como fuera compatible con el nivel de bienestar material, teniendo en cuenta la tecnología accesible a los nativos. En ese caso la incursión de los extranjeros debe por sí misma reducir el nivel existente de bienestar o disminuir la población nativa y reducir también las expectativas futuras por lo menos de algunos de los habitantes, de modo tal que una incursión semejante, aun cuando sólo estuviera destinada a compartir los

pastos comunes sin una oposición justificable de los nativos, ha de ser en sí misma injustificable, ha de constituir una pretensión de derechos que están más allá de lo que autoriza la estipulación.

Pero los europeos llevaron consigo una tecnología superior que le permitió a una mayor cantidad de gente vivir más tiempo y más holgadamente, con más oportunidades (aunque algo diferentes) de las que antes disponía. Si se les hubiese brindado a los indios y a los inuit una posibilidad efectiva de adaptar sus estilos de vida a fin de compartir ese mejoramiento general y se les hubiese dejado la libertad de optar por sus antiguos estilos de vida aunque con sus expectativas reducidas, de modo tal que pudieran elegir, no podría decirse que la situación de los nativos empeoró y éstos no hubieran podido quejarse con justicia de la incursión de los europeos ni de que éstos se apropiaran de gran parte del continente, aunque evidentemente no de todo.

La interpretación más precipitada de la historia de la apropiación que los europeos hicieron de Norteamérica muestra que no se les ofreció a los pobladores nativos una elección efectiva entre adaptarse a los modos europeos y conservar sus propios modos. Las reservaciones fueron proyectadas para permitir que se conservaran los antiguos estilos de vida, pero a los indios les estaba efectivamente negada la incorporación a la sociedad de los hombres blancos. Por consiguiente, la condición de los nativos empeoró y se violó la estipulación. Pero esa violación fue un efecto innecesario, aunque no sorprendente, de una invasión basada en una tecnología que ofreció una forma potencialmente justificable de adquisición. De ahí que hubiera sido posible que los habitantes no nativos de Norteamérica les ofrecieran a los nativos una negociación o un acuerdo basado en una posición inicial justa sin abandonar el volumen de su apropiación.

La apropiación europea —como cualquier adquisición en el estado de naturaleza— no se basó en el consentimiento. Una negociación justa debe implicar en sí misma el consentimiento o, por lo menos, un consentimiento hipotético fundamentado racionalmente; pero la apropiación inicial no es una negociación. Se da en condiciones en las cuales cada actor es libre de perseguir sus propios intereses salvo si al hacerlo saca ventaja

de algún otro actor. Exigir el consentimiento en los actos de apropiación que establecen derechos iniciales sería una forma de fomentar la aparición de los jugadores líberos. Sería imponer una barrera, con toda probabilidad insuperable, al progreso humano, puesto que semejante progreso sólo puede lograrse mediante cambios radicales de los estilos de vida. Los pueblos nativos de Norteamérica no pueden quejarse, con razón o con justicia, de que los europeos hayan sido incapaces de ganarse el consentimiento previo de los pobladores originales antes de establecerse en el nuevo mundo. Sin duda habría sido imposible obtener semejante consentimiento; los indios y los inuit probablemente no podían comprender previamente esa nueva tecnología que si bien les daría nuevas oportunidades, suplantaba sus antiguos estilos de vida.

Esos antiguos estilos de vida estaban efectivamente condenados a desaparecer a causa de la llegada de los europeos. Eran estilos en alto grado adaptados a un ambiente que los pobladores originales no podían transformar. Pero la tecnología europea hizo que esa adaptación resultara anticuada. Esa tecnología resolvió los problemas contra los que luchaban los nativos y seguramente trajo otros nuevos y diferentes que, sin embargo, estaban relacionados con un nivel más avanzado de desarrollo humano. Las formas de producción y las estructuras de creencia que habían sido buenas para los habitantes originales de América perdieron toda significación a la luz de una tecnología superior y una ciencia más general. El efecto que tuvo la invasión europea fue convertir en una forma de juego las prácticas culturales que hasta entonces habían sido necesarias para la supervivencia. Porque los europeos ofrecieron un estilo de vida que podía mantener a más gente, con expectativas de vida más larga y un nivel de bienestar material más alto y porque ofrecieron además un horizonte más amplio de oportunidades.

Decir que la tecnología europea hizo que la cultura de los indios y los inuit resultara anticuada no implica denigrar esa cultura ni despreciar los logros alcanzados por esos individuos. La obsolescencia de esas culturas sólo muestra que el progreso humano continúa siguiendo los criterios enumerados en 5.1. Aquellos que condenan el hecho de que una cultura suplante a

otra y lo consideran un "genocidio cultural" aparentemente prefieren que los seres humanos sean piezas de museo vivas y que participen de una interminable reformulación de los roles del pasado, en lugar de dejarlos en libertad de participar en el curso del desarrollo humano. La estipulación, al permitir aquellas apropiaciones que ponen en evidencia la obsolescencia de ciertos estilos de vida, estimula a los individuos que se adhieren a ella a aprovechar las nuevas oportunidades que han de permitirles participar otra vez en ese desarrollo. Desafortunadamente, e injustificadamente, los europeos prefirieron con excesiva frecuencia mostrarles a los indios y a los inuit la obsolescencia de los estilos de vida que llevaban, sin ofrecerles la oportunidad de participar en el estilo de vida que suplantaba a aquéllos.

Hasta ahora hemos ignorado una dimensión adicional de la apropiación hecha por los europeos en América del Norte. Nos referimos al efecto que tuvo esa invasión no en los pobladores nativos, sino en los demás pueblos de la tierra. En el corto plazo, esos efectos fueron mínimos, pero como vimos, las consideraciones de corto plazo no bastan para fundamentar una adquisición permanente. Si se supone que la presencia de los europeos en América del Norte excluyó a los africanos y a los asiáticos de este continente y sólo les permitió la entrada como inmigrantes en términos aceptables para los propietarios, ¿no implicó acaso esa actitud una restricción de las oportunidades para esos otros pueblos?

Los africanos y los asiáticos habrían mejorado su situación si hubiesen desarrollado la tecnología que les permitiera llegar primero a América y apropiarse de ella. Pero la incapacidad de actuar primero no empeora en sí misma la situación de alguien. Debemos preguntarnos si los africanos y los asiáticos se habrían beneficiado en el caso de que América hubiese quedado sencillamente en posesión de los indios y los inuit. Los hechos pueden resultar poco claros. Pero, seguramente es plausible sostener que la invasión europea en América constituyó una parte integrante de una enorme expansión de la capacidad productiva global que efectivamente amplió los horizontes, mejoró las condiciones de vida, alargó las expectativas de vida y expandió las poblaciones a lo largo y a lo ancho del mundo. El hecho de que

esos beneficios se hayan distribuido de manera no equitativa no es en sí mismo una prueba de injusticia. Y lo que es más serio, algunos pueblos compartieron, aunque escasamente, las mejoras. Pero ni siquiera esto es una señal de injusticia, ya que es posible que esos pueblos prefirieran conservar su estilo de vida a pesar de los costos que ello les provocaba. Pero seguramente muy pocos creen que a todos los pueblos se les brindó una oportunidad efectiva de participar justamente del aumento del bienestar global y del progreso humano general alcanzado gracias a la tecnología europea. No obstante, las verdaderas y extensas injusticias que se cometieron podrían parecer incidentales en relación con nuestro interés particular; podrían haberse evitado sin invalidar por ello el acto de apropiación de los europeos o, mejor dicho, sin obstaculizar los demás acontecimientos que multiplicaron la capacidad productiva humana.

Al considerar el acto de apropiación, por parte de miembros de una sociedad más avanzada, de tierras antes habitadas sólo por miembros de sociedades menos avanzadas, nos hemos internado en una cuestión de gran complejidad. No pretendemos que nuestro análisis sea conclusivo, sino simplemente mostrar la relevancia que adquiere la estipulación cuando se plantean problemas de apropiación entre los pueblos, y señalar el tipo de conclusiones a las que puede conducir su aplicación. Y sobre todo, queremos rescatar de la difamación indiscriminada que se ha extendido sobre la idea imperial, esa fuerza completamente real y significativa por la cual una potencia más avanzada trata de mejorar su propia situación de un modo que permite el acceso efectivo de otros pueblos a una perspectiva de aumentar la población, prolongar sus vidas, acrecentar sus bienes materiales y multiplicar sus oportunidades.

6.1 Las relaciones entre las personas de las diferentes generaciones parecen escapar al alcance de la moral racional establecido por nuestra teoría. Si bien supusimos un desinterés mutuo, al apelar a los beneficios de la cooperación establecimos restricciones imparciales a la maximización de la utilidad individual. Pero la cooperación sólo es posible entre personas contemporáneas que realmente están en interacción. Aunque un individuo pueda hacer mucho por beneficiar o por dañar a sus descendientes, sólo aquellos cuyas vidas coinciden en el tiempo

con la de tal individuo pueden a su vez beneficiarlo o dañarlo. Por lo tanto, él no tiene razones para interesarse por los demás. Al maximizar su propia utilidad sin preocuparse por las personas futuras cuyas vidas no han de coincidir en el tiempo con la propia, nuestro individuo no corre ningún riesgo de llegar a un resultado subóptimo. Ninguna restricción que el individuo aceptara podría dar como resultado una simple transferencia de sus utilidades a personas futuras, ni podría conducir a la obtención de beneficios mutuos. Todo costo de sus acciones que el individuo les imponga a las personas futuras es estrictamente incidental respecto de los beneficios que recibe. Por lo tanto, no puede constituir una violación a la estipulación. Por consiguiente, un individuo no comete ninguna injusticia contra sus descendientes al no preocuparse por ellos. Si no les deja un mundo que pueda ser adecuadamente habitado por ellos y por lo tanto no les deja un bienestar, esto sólo ha de caracterizar las circunstancias en las cuales les tocará vivir a esas personas futuras; de modo que no se puede decir que se afecten los derechos de tales personas.

Si ésta es realmente la posición en que nos coloca la razón, debemos pues aceptarla admitiendo otra diferencia, muy profunda, entre nuestra moral intuitiva y la que armoniza con la elección racional. Estaríamos pues obligados a suponer que las relaciones morales entre las personas de diferentes generaciones exigen una base afectiva. Al rechazar el supuesto de mutuo desinterés, deberíamos indagar la moral de los sentimientos. Lo normal es que las personas se interesen por lo menos en sus descendientes inmediatos. Y ese interés queda expresado en el deseo de convertir al mundo en un sitio mejor, o por lo menos no en un lugar peor que el que encontramos nosotros. Las comunidades humanas se conservan a través del tiempo; y las personas consideran que sus descendientes son miembros de esa sociedad como ellos mismos lo son. Al pensar de ese modo, los individuos consideran los efectos adversos que puedan tener sus actividades en los miembros futuros de su sociedad como costos realmente desplazados, que violan el requisito basado en la estipulación que establece que los costos deben internalizarse. De modo tal que deberíamos introducir la idea de que los seres humanos están comprometidos en una empresa conjunta que se extiende

a través del tiempo destinada a crear y a mantener una sociedad: un estilo de vida del que todos participan. Deberíamos llegar a la concepción de sociedad sostenida por Burke, según la cual se trata de una empresa basada en un contrato establecido entre las generaciones pasadas, presentes y futuras.[16] Con lo cual podríamos aplicar nuestra teoría moral a fin de determinar una estructura de derechos y obligaciones entre los miembros de las diferentes generaciones, basada en la ficción de un acuerdo, aun cuando no haya interacción.

Pero, ¿debemos partir necesariamente del supuesto de mutuo desinterés? El argumento de que un individuo ignora a aquellos descendientes cuyas vidas no han de coincidir en el tiempo con la suya propia, ¿no está acaso ignorando la verdadera significación de la superposición de las generaciones? Las generaciones de la humanidad no aparecen sobre el escenario de la vida y se retiran de él, de a una generación por vez. Cada persona entra en interacción con otras, de más y de menos edad que ella, y por lo tanto forma parte de un hilo continuo de interacción que se extiende desde el pasado humano más remoto hasta el futuro más lejano de nuestra especie. La cooperación mutuamente beneficiosa incluye directamente a personas de diferentes generaciones cuyas vidas coinciden en el tiempo, pero esto crea lazos cooperativos indirectos que se extienden a lo largo de la historia. Cada persona, al considerar los términos en los cuales ha de cooperar con aquellos de una generación anterior a la suya, debe tener en cuenta la necesidad de establecer términos similares con aquellos de la generación posterior a la suya, quienes a su vez deben tener en cuanta la necesidad de cooperar con los miembros de la próxima generación, y así sucesivamente. Por lo tanto, aunque cada individuo esté dispuesto a convenir con sus contemporáneos que deberían agotar los recursos del planeta sin pensar en aquellos que aún no nacieron, la necesidad de continuar cualquier acuerdo a lo largo del tiempo, la necesidad de extender dicho acuerdo a aquellos que nacen a medida que lo abandonan quienes mueren, garantiza que, entre las personas racionales, los términos del convenio deben permanecer constantes, de modo tal que el agotamiento de los recursos del planeta no es una opción aceptable. Independientemente del momento en que viva, el

individuo debería esperar los mismos beneficios relativos de la interacción con sus semejantes que recibieron sus predecesores y que, a su vez, han de recibir sus sucesores.

Partamos pues de nuestro supuesto de desinterés mutuo y preguntémonos qué condiciones rigen la transferencia de los derechos de jurisdicción y de propiedad que hacen las personas a sus sucesores temporales; y preguntémonos también qué condiciones rigen los términos de los acuerdos cooperativos en los cuales se espera que los participantes sean reemplazados por sucesores. Específicamente, analicemos las condiciones que gobiernan la herencia y la inversión.

6.2 El supuesto de desinterés mutuo parece anular los problemas de herencia.[17] Una persona que literalmente no se interesa por sus semejantes ni por los intereses de éstos, debe ser completamente indiferente en cuanto a lo que ocurra con sus posesiones y sus privilegios cuando ella ya no pueda obtener beneficios de todo eso. Durante su vida esa persona dispone de tales posesiones del modo que maximice su utilidad presente. Establece acuerdos con otras personas mediante los cuales intercambia derechos de los que puede hacer uso y disfrutar en vida por derechos que les cederá a esas otras personas a su muerte. Si la persona no dispusiera así de todas sus posesiones y privilegios, a su muerte esos bienes quedarían sin dueño y otros podrían apropiarse de ellos. Puesto que la existencia de semejantes bienes sin dueño probablemente provocaría inconvenientes en el seno de la sociedad, podemos suponer que ha de establecerse algún procedimiento que disponga de ellos, un procedimiento que, si satisface el principio de concesión relativa minimáxima, ha de proporcionarle a cada miembro de la sociedad una porción equivalente del beneficio esperado que produzcan esos bienes. La herencia como tal no existiría, pues no es una práctica que las personas mutuamente desinteresadas instituirían o aceptarían racionalmente.

Enfoquemos ahora la herencia desde un punto de vista más realista. Suponemos, como es claramente el caso, que a la mayor parte de las personas le interesa cómo disponer de sus posesiones y sus privilegios; que, en realidad, la mayor parte desea compartirlos con su cónyuge y legarlos a sus descendientes. El desinterés mutuo prevalece, no entre las partes de la relación

constituida por el legado y la herencia, sino entre los miembros de la sociedad para quienes es necesario definir una práctica aceptable de herencia, de modo tal que cada uno pueda reconocer la transferencia de derechos que implique tal práctica. ¿Qué derechos deben reconocer racionalmente las personas que no se interesan ni por la creación individual de un legado ni por la recepción individual de una herencia?

Pero, ¿no respondimos ya esa pregunta cuando explicamos qué era la apropiación legítima? Si un individuo adquiere un derecho al uso exclusivo de cierta porción de tierra, de algún otro bien, o un derecho al ejercicio exclusivo de algún privilegio, con seguridad ese individuo está autorizado a disponer de ese derecho como mejor le plazca y por consiguiente a legárselo a otra persona al morir. Pero el asunto no es tan sencillo. Hemos dicho que un derecho adquirido es permanente, pero no hemos defendido esa afirmación. Supusimos que los derechos son transferibles, pero, ¿no puede haber ciertos límites en las formas admisibles de transferencia? Los derechos de un individuo constituyen la dote que éste aporta a la interacción cooperativa y de mercado. Por lo tanto está autorizado a intercambiar sus derechos en las prácticas cooperativas o de mercado. Lo cual no significa que esté autorizado a disponer de sus derechos de modos diferentes, por ejemplo, regalándolos o dejándolos como herencia. Quizá nadie esté autorizado a adquirir un derecho, excepto mediante la apropiación inicial cuando no existe un propietario anterior de tal derecho, o mediante una transacción de mercado en la cual el propietario anterior recibe un pago o una compensación. Podríamos imaginar entonces una sociedad en la cual los derechos expiraran con su propietario y sólo pudiera disponerse de ellos a través del intercambio de mercado, o mediante el abandono. En este último caso, la posesión o el privilegio sería reasignado mediante algún procedimiento social que satisficiera la concesión relativa minimáxima. Una vez más la herencia no existe.

Sin embargo, no suponemos que las personas racionales tengan que rechazar los derechos permanentes de los cuales pueden disponer no sólo mediante el intercambio de mercado, sino también regalándolos o legándolos, según los deseos del poseedor. Porque es evidente que el deseo de adquirir tierra u

otros bienes que luego puedan transferirse a los descendientes ha sido una poderosa motivación para las actividades productivas que aumentó el bienestar material y amplió la cantidad de oportunidades individuales. La persona que reclama un derecho permanente y transferible sobre cierto bien, puede beneficiarse luego invocando una práctica que redunde en la ventaja esperada de todos. El argumento que ofrecemos a fin de defender la apropiación individual de la tierra y de otros bienes en 7.4.1 puede extenderse y fundamentar un derecho de legado.

El derecho a disponer de las propias posesiones regalándolas o legándolas no es una parte integrante del derecho al uso exclusivo de tales posesiones como lo exigen la interacción de mercado y la interacción cooperativa. Su justificación racional depende del interés que tenga el individuo por cómo ha de disponer de sus bienes, lo cual supone un interés por otras personas particulares. Y esta suposición es ajena al mercado e innecesaria para la cooperación. Pero la racionalidad del derecho a disponer de los propios bienes se refiere a personas que no se interesan mutuamente unas por las otras. Le reconozco a alguien el derecho a legar sus posesiones como mejor le plazca, aun cuando no me interese por esa persona, porque la práctica del legado satisface la estipulación, ya que les permite a aquellos que lo reciben beneficiarse sin sacar de ningún modo ventaja de sus semejantes.

Pero para que el legado no esté en conflicto con la estipulación lockeana, debe respetar ciertos límites. Porque la herencia de cada persona constituye un componente significativo de sus recursos disponibles iniciales y el volumen de los recursos de un individuo afecta no sólo las oportunidades de que éste dispone, sino también las oportunidades accesibles a los demás. Imaginemos una comunidad de pequeños terratenientes, en la cual todos se benefician con el estímulo a la producción que producen las propiedades privadas, con la especialización limitada y con los intercambios que resultan de esas posesiones privadas. Supongamos que, por la razón que fuere, todos los terratenientes deciden entregarle su propiedad a un solo hombre, cuya personalidad carismática (y la promesa de que serán recompensados en el cielo) ha encandilado a todos. Esa herencia puede convertir a dicho hombre en el propietario de un monopolio sobre

la tierra y a los demás en trabajadores desposeídos. Un derecho a una posesión tan extensa —toda la tierra cultivable de la comunidad—parecería violar claramente la estipulación. El derecho al legado no puede ser tan amplio que le proporcione a un individuo cualquiera una herencia que éste no podría haber adquirido de ningún otro modo sin restringir las oportunidades o reducir el bienestar de sus semejantes.

Los límites aceptables de la herencia no están sujetos a una determinación a priori. Si el derecho al legado no constituye un estímulo a la producción que se suma a los beneficios esperados de todos los miembros de la comunidad, podemos suponer que esas herencias se han de circunscribir más estrechamente que si el deseo de asegurar el porvenir de los propios descendientes constituye una fuerte motivación para que el individuo invierta su tiempo y su esfuerzo. Aquí debemos contentarnos con llegar a estas conclusiones generales que indican cómo, siguiendo nuestra teoría, los problemas relativos a los derechos de legado y herencia quedan resueltos.

6.3 Nuestra teoría, que se basa en la representación de la sociedad como una negociación o un acuerdo en desarrollo establecido entre personas mutuamente desinteresadas da una solución muy clara al problema de la tasa racional de inversión —ahorros de la producción presente para aumentar la producción futura—.[18] Uno de los componentes fundamentales del progreso social, como lo dijimos en 5.1 es el aumento de la oferta per capita de los bienes materiales. Suponiendo que esa inversión sea productiva, de modo tal que haya una tasa positiva de acumulación antes que una tasa negativa o una tasa positiva de agotamiento, ese progreso se logra invirtiendo una porción de la producción presente como capital para la producción futura en lugar de utilizarlo para el consumo presente. Pero si con esto se logra aumentar la oferta futura de bienes de consumo, ello no se obtiene disminuyendo la oferta presente. Aquellos que ahorran para aportar capital a la producción futura deberían obrar mejor, atendiendo a sus propias utilidades, que aquellos que en cambio consumen todo lo producido. Puesto que el ahorro permite que haya más bienes en el futuro de los que podrían obtenerse en el presente, constituye una transferencia de aquellos que poseen menos bienes a aquellos que poseen más. Así

presentada la cuestión, semejante transferencia parecería violar el principio de concesión relativa minimáxima. De ahí que nuestra teoría parecería exigir una tasa cero de inversión y por lo tanto estaría negando la racionalidad de elegir el progreso social.

Pero en verdad no es así. Consideremos los términos de una negociación racional entre personas de diferentes generaciones. Puede suponer que la posición negociadora inicial es la misma, sea cual fuera la generación a la que uno pertenece. Porque esta posición está determinada por la interacción en estado de naturaleza restringida por la estipulación. Y podemos suponer que el estado de naturaleza brinda expectativas en general constantes en el tiempo a todos los individuos, teniendo en cuenta que los miembros de cada generación están protegidos por la estipulación de las actividades de sus predecesores que de otro modo podrían agotar los recursos disponibles y cercenar sus oportunidades. Por consiguiente podemos decir que las diferencias materiales individuales de la posición negociadora inicial reflejan factores que no dependen del tiempo. La posición inicial del individuo promedio es la misma, sea cual fuere la generación a la que pertenezca.

Pero la pretensión máxima a la que puede racionalmente aspirar cada individuo depende del tiempo. La pretensión máxima de cada persona es el mayor rendimiento posible que puede obtener de la cooperación, suponiendo que los demás reciban por lo menos lo que podrían esperar de la interacción en estado de naturaleza restringida por la estipulación. Por lo tanto, si el individuo es un miembro de la generación inicial, su pretensión se basa únicamente en los frutos de la interacción social mantenida entre los miembros de esa generación. Si el individuo es un miembro de la segunda generación, su pretensión se basa en los frutos de la interacción social mantenida entre los miembros de su propia generación y también en los frutos de la inversión realizada por la generación inicial, de todo el superávit cooperativo, es decir de aquello que pudieron producir por encima de lo que podrían haber esperado obtener en ausencia de la cooperación. Si un individuo es un miembro de la tercera generación, basa su pretensión en los frutos de la interacción mantenida entre los miembros de su propia generación y también en los frutos de la inversión realizada por las dos primeras generaciones, de todo el superávit cooperativo. Y así sucesivamente, de modo tal que, en la medida en que la inversión sea productiva, los miembros de cada generación puedan aspirar a pretensiones máximas mayores que los miembros de las generaciones precedentes.

Aplicando el principio de concesión relativa minimáxima a las pretensiones máximas y suponiendo el carácter invariable de la posición negociadora inicial, comprobamos por supuesto que los miembros de cada generación reciben más, en términos absolutos, que lo que recibieron sus antecesores. En una negociación entre generaciones, no hay un bien único transferible que pueda ser distribuido entre todos, sino que hay un bien en firme crecimiento que se distribuye en cantidades igualmente proporcionales. Los miembros de cada generación reciben la misma proporción de sus pretensiones que la que recibieron quienes los precedieron y la que recibirán los miembros futuros de esa sociedad, pero la productividad de la inversión garantiza el enriquecimiento continuo.

Por consiguiente, la tasa racional de inversión se determina aplicando la concesión relativa minimáxima a las pretensiones basadas en la tasa de acumulación posible. Podemos comparar esta solución con la que ofrecen el principio de bienestar absoluto maximínimo y el principio utilitarista de bienestar promedio máximo. Aunque en muchas circunstancias esos principios alternativos conduzcan a prescripciones idénticas, sus aplicaciones difieren notablemente y nuestra posición se sitúa entre ambos.

A fin de maximizar el nivel mínimo de bienestar absoluto debería prohibirse cualquier transferencia de aquellos que poseen menos a aquellos que poseen más. De ahí que debería prohibirse que la generación inicial haga cualquier ahorro; sus miembros deberían consumir todos los bienes que producen. Pero entonces cada generación estaría en la posición de la generación inicial: cada una debería consumir todo lo que produce, puesto que si no lo hiciera tendría menos y así sus sucesores podrían tener más. O antes bien, cada generación debería ahorrar sólo lo necesario para contrarrestar cualquier agotamiento de los recursos, puesto que cada generación debe

dejar a sus sucesores por los menos tan bien situados como ella misma. Al maximizar el nivel mínimo de bienestar absoluto deberíamos invocar la estipulación, pero entonces no habría cooperación intergeneracional y la tasa racional de inversión sería normalmente igual a cero.[19]

Por otra parte, teniendo en cuenta los supuestos plausibles sobre los efectos productivos de la inversión, maximizar el bienestar promedio a través del tiempo exigiría una tasa muy alta de inversión. Porque aun teniendo en cuenta el supuesto habitual de que la utilidad derivada de cada unidad de un bien disminuye a medida que aumenta la riqueza total del individuo, una tasa suficientemente alta de compensación sobre la inversión traducida en bienes proporciona una tasa positiva menor pero estable de compensación traducida en utilidades. Puesto que los individuos no están en sí mismos habilitados para disponer así de sus bienes, el utilitarista puede encontrarse en la posición de defender la restricción del consumo corriente de todo aquello que sea necesario para asegurar la productividad máxima y la renuncia a todo lujo presente, ya que en sus cálculos los mayores beneficios eventuales que han de recibir las personas futuras exceden los costos que recaen sobre las personas actuales.

Entre estos extremos, nuestra teoría, al igualar los beneficios relativos antes que los absolutos, defiende el progreso sin ignorar por ello las pretensiones de las personas actuales. Al considerar que la sociedad es una negociación en la cual los términos permanecen constantes a través del tiempo, de modo tal que cada generación le ofrece a su sucesora el mismo acuerdo que ella misma aceptó de sus predecesores, podemos dar una clara explicación de la justicia entre las diferentes generaciones.

Notas

1 Véase la nota 26 del capítulo 1, *supra*.

2 *Republic*, 357a-367e.

3 El anunciante es la Miller Brewing Company

4 Sobre un punto de vista opuesto, véase T. Nagel, *Moral Questions*, Cambridge, 1979, pág. X. "Creo que uno debería confiar más en los problemas que en las soluciones, en la intuición que en los

argumentos en el desacuerdo pluralista que en la armonía sistemática. La simplicidad y la elegancia nunca han tenido razones para pensar que una teoría filosófica es cierta: por el contrario, generalmente tienen fundamentos para creerla falsa". Pero, ¿por qué debería la filosofía diferir de ese modo de la ciencia? Con todo, no hay una explicación mejor del punto de vista opuesto a aquel en el que se fundamenta *Una moral basada en el acuerdo*.

5 Véase J. Rawls, *A Theory of Justice*, Cambridge, Mass. 1971, pág. 20. También N. Daniels: "Wide Reflective Equilibrium and Theory Acceptance in Ethics", *Journal of Philosophy* 76, 1979, págs. 256-282.

6 Véase la discusión de los bienes puramente públicos en H.van den Doel, *Democracy and Welfare Economics*, traducción de B. Biggins, Cambridge, 1979, págs. 25-30.

7 Véase M. Friedman, *Capitalism and Freedom*, Chicago, 1962, págs. 174-176.

8 Véase R. Nozick, *Anarchy, State, and Utopia*, Nueva York, 1974, págs. 160-163.

9 *Ibid.*, pág. 161.

10 Rawls, *op.cit.*, págs. 101-102.

11 *Ibid.*, pág.62

12 *Ibid.*, pág.101.

13 *Ibid.*, pág. 250.

14 Véase N. Eberstadt, "The Health Crisis in the USSR", *New York Review of Books* 28, n°2, 1981, pág. 23. El autor acepta para la duración promedio de la vida a final del siglo una cifra estimativa más cercana a los 30 que a los 40 años. Pero en otra parte yo he leído cifras estimadas más cercanas a los 40.

15 Sobre una discusión de este problema, véase M. McDonald, "Aboriginal Rights", en W. R. Shea y J. King-Farlow (comps.), *Contemporary Issues in Political Philosophy*, Nueva York, 1976, págs. 27-48. McDonald sostiene que "este país (Canadá) debería reconocer los derechos de los aborígenes *sobre la base de la adquisición original*", pág.47.

16 Véase Edmund Burke, *Reflections on The Revolution in France* (Londres, 1970); introducción A. J. Grieve, Londres, 1910, pág. 93.

17 Sobre una discusión anterior de la herencia, relevante pero parcialmente invalidada por el presente argumento, véase mi "Economic Racionality and Moral Constraints", *Midwest Studies in Philosophy* 3, 1978, págs. 84-88.

18 Sobre algunas discusiones previas, véase mi "The Social Contract: Individual Decision or Collective Bargain?", *Philosophy and Public Affairs* 6, 1977, págs. 63-65, y "Social Choice and Distributive Justice", *Philosophia* 7, 1978, págs. 249-250.

19 Para evitar esa conclusión, Rawls ajusta su enfoque de la motivación considerando a las personas como "representantes de líneas familiares, digamos, con vínculos de sentimiento entre las sucesivas generaciones", *A Theory of Justice*, pág. 292.

X

El anillo de Giges

1.1 La leyenda es familiar.[1] Se decía que Giges, rey de Lidia, poseía un anillo mágico: si su portador le hacía girar el engaste lograba hacerse invisible. El padre de Giges había sido un pastor al servicio del entonces rey. Un día, mientras el padre de Giges vigilaba su rebaño, se levantó una gran tormenta; la tierra se abrió y el pastor descendió a una caverna en la que vio un caballo de bronce con una abertura en el cuerpo dentro de la cual había un gigantesco cadáver que llevaba puesto un anillo de oro. El pastor se apoderó del anillo y luego, reunido con los demás pastores, hizo girar por casualidad el engarce, con lo cual se volvió invisible a los ojos de éstos. El padre de Giges hizo girar nuevamente el engarce y volvió a ser visible. Consciente del poder que había adquirido gracias al anillo, el pastor se las ingenió para convertirse en uno de los mensajeros que informaba al rey sobre los rebaños. Gracias a la magia que lo hacía invisible a voluntad, se las arregló para llegar hasta el lecho de la reina. La satisfacción que obtuvo ésta de ese encuentro la indujo a enviar al pastor a la cámara del rey con un recado por completo diferente. Y de ese modo, el pastor obtuvo para sí y sus descendientes el trono de Lidia.

Glaucón nos señala la moraleja de este relato, al exponer la tesis de que la injusticia es superior a la justicia, tesis que Sócrates debía refutar.

> Ahora bien, si existieran dos anillos como ése, y el hombre justo se colocara uno y el hombre injusto el otro, parecería que nadie sería tan firme para ajustarse a la justicia y mantenerse alejado de aquello que les pertenece a otros y no apoderarse de ello, aun cuando tuviera la oportunidad de tomar sin temor lo que desee del mercado, de introducirse en las casas y mantener comercio carnal con quien quiera, de matar o quitar las cadenas a quien desee, y de hacer otras cosas como si fuera un dios entre los seres humanos. Y al hacerlo, ambos obrarían del mismo modo, ambos recorrerían el mismo camino.[2]

El anillo coloca a su portador fuera de las circunstancias de la justicia. Y esas son las circunstancias de la debilidad. Como sostiene Glaucón antes de relatar la leyenda,

> Dicen que cometer injusticia es naturalmente bueno y que sufrir injusticia es malo, pero que el mal que implica sufrir una injusticia excede ampliamente el bien que implica cometerla; de modo tal que, cuando dos personas cometen injusticias recíprocas y las sufren y prueban el sabor que producen, parece conveniente —para aquellos que no pueden rehuir una y elegir la otra— establecer un pacto entre ellos que determine que no han de cometerse ni sufrirse injusticias... Y esto... es el origen y el ser de la justicia; es un punto medio situado entre lo mejor —cometer injusticias sin pagar por ello— y lo peor —sufrir una injusticia sin tener la posibilidad de vengarse—. Lo justo está en el medio de esas dos situaciones, aceptado no porque sea bueno, sino porque se lo acata debido a la falta de vigor para cometer injusticias. El hombre capaz de cometerlas y que es un verdadero hombre jamás establecería un pacto con nadie para no cometer ni sufrir injusticia. Para hacerlo tendría que estar loco.[3]

El problema de Glaucón no es el de nuestro familiar Dilema del Prisionero. En realidad, Glaucón supone que las personas pactan en favor de la justicia como una respuesta a situaciones del tipo del Dilema del Prisionero, en las cuales cada individuo perdería más a manos de los otros de lo que ganaría a sus expensas. Pero Glaucón afirma que una persona tiene razones para acatar las restricciones que impone la justicia sólo a causa de su debilidad, que se manifiesta en su incapacidad de sacar una ventaja neta de sus semejantes. La persona verdaderamente fuerte rechazaría no sólo el acatamiento sino la cooperación misma, ignorando tanto la tercera ley de la naturaleza de Hobbes (que los hombres cumplan sus pactos), como la segunda (que cada individuo renuncie a una porción de su derecho inicialmente ilimitado de naturaleza, siempre que los demás también lo hagan). Al poseer el anillo, el pastor no ve que establecer acuerdos con sus semejantes le proporcione ningún beneficio.

Puede parecer que lo que dice Glaucón sobre la justicia es lo que debemos decir nosotros sobre la moral. O sea, que es un punto medio situado entre lo mejor —mejorar nuestra situación, sea cual fuere el costo que esto les depare a los demás— y

lo peor —que los demás empeoren nuestra situación como les plazca—. Nos importa la moral, no por la moral misma, sino porque carecemos de las fuerzas necesarias para dominar a nuestros semejantes o de la autosuficiencia para prescindir de la interacción con ellos. La persona que puede alcanzar sus fines obrando independientemente de los demás o subordinándolos, nunca acataría las restricciones de la moral. Para hacerlo tendría que ser irracional, loca.

Tanto el argumento de Hobbes como el de Hume parecerían confirmar que una moral racional es un mal necesario. Para Hobbes, las leyes de la naturaleza, "la única y verdadera Filosofía Moral", como las llama, constituyen la base de toda restricción;[4] la segunda ley nos dice que hay que renunciar a una porción de nuestro derecho natural ilimitado inicial. Pero, como señala el propio Hobbes, "Como todos los hombres tenían por naturaleza derecho a todas las cosas, cada uno tendría derecho a reinar sobre todos los demás. Pero...esto afectaría la seguridad de cada individuo, y ese derecho capacitaría a los hombres, en virtud de un consentimiento común, a gobernarse y defenderse; mientras hubiera un hombre de poder irresistible, no habría ninguna razón por la que no debiera ser ese poder quien se gobernara y defendiera a sí mismo y gobernara y defendiera a los otros hombres de conformidad con su propia discreción".[5] Según afirma Hobbes, ésta es la base de la soberanía que ejerce Dios sobre los hombres —un punto que vale la pena tener presente para relacionarlo con la idea de que el anillo de Giges confiere el poder de "hacer otras cosas como si fuera un dios entre los seres humanos". Pero lo que nos interesa en este momento son las implicaciones que tiene el argumento de Hobbes en la moral. Como un tipo de restricción, queda claro que tal argumento no constituye una apelación racional para aquellos que poseen el poder suficiente para dominar a sus semejantes; tales personas no tienen ninguna razón para renunciar a una porción de sus derechos permisivos ilimitados ni renunciar a abrazar la maximización directa.

Aunque Hume basa la moral en los sentimientos de simpatía, y por lo tanto tuísticos, de las personas, su enfoque de la justicia entra dentro de la tradición contractual.[6] La obligación primaria que tiene cada persona para con la justicia se funda-

menta en el interés; si no existe tal interés, la justicia no tiene cabida. Hume observa que

> Si existieran especies de criaturas entremezcladas con los hombres las cuales, siendo racionales poseyeran sin embargo fuerzas inferiores a las de los seres humanos, tanto físicas como mentales, es decir que fueran incapaces de ofrecer alguna resistencia y no pudieran nunca, ante las mayores provocaciones, hacernos sentir los efectos de su resentimiento; la consecuencia necesaria sería que...nosotros, para decirlo con propiedad, no deberíamos acatar ninguna restricción de la justicia respecto de tales criaturas... Las relaciones que mantuviéramos con ellas no podrían llamarse sociedad, porque esta palabra supone cierto grado de igualdad; pero, ejercido el dominio absoluto por un lado y manifestada la obediencia servil por el otro... como el ejercicio del poder, establecido tan firmemente en la naturaleza, no provocaría nunca un inconveniente, las restricciones de la justicia y la propiedad, al ser por completo *inútiles*, nunca tendrían un lugar en una confederación tan desigual.[7]

Además Hume señala que la justicia depende de la incapacidad individual, de la necesidad que cada persona tiene de las demás.

> Si las especies humanas estuvieran así formadas por la naturaleza, como cada individuo posee en sí mismo todas las facultades necesarias tanto para su propia preservación como para la propagación de su especie... parece evidente que un ser tan solitario sería tan incapaz de obrar con justicia como de participar del discurso y la conversación sociales. Si el interés mutuo y la indulgencia no sirvieran a ningún propósito, nunca podrían dictar la conducta de ningún hombre razonable... Y como aquí suponemos que cada hombre sólo se ama a sí mismo, y que para lograr su seguridad y su felicidad sólo depende de sí mismo y de su propia actividad, el individuo debería en toda ocasión... imponer su preferencia a la de cualquier otro, pues no se halla ligado a ninguno de ellos por ningún vínculo ni de naturaleza ni de interés.[8]

Hume rechaza cualquier concepción de la racionalidad práctica. Pero allí donde él habla de interés, nosotros podemos hablar de razón. De modo que podemos reinterpretar a Hume y suponer que él retoma la tesis de Glaucón y la desarrolla más

detalladamente, afirmando que una moral racional exige una igualdad relativa y la incapacidad individual de aquellos compelidos por ella. Dependemos de aquellas otras personas a las que no podemos dominar; de modo que debemos adaptarnos a ellas. Pero al advertir esto, cobramos también conciencia de que somos menos de lo que deseamos ser. Como un remedio para la debilidad humana (y también como un recordatorio de ella), la moral no puede tener una influencia más profunda en los seres humanos. Si esta desagradable conclusión no aparece en Hume, ello sólo se debe a que Hume acepta una moral afectiva, basada en la transmisión simpática de los sentimientos de una persona a otra. Pero una moral semejante no tiene ninguna base en la racionalidad maximizadora de la utilidad.

La moral considerada como un mal necesario, seguramente no ha de despertar la reverencia que, según se piensa, le profesaba Kant, pero, ¿ha de cumplir siquiera con su cometido?[9] Al procurar que Socrates lo persuada de que la justicia debería estimarse por lo que es en sí misma, ¿no estaba acaso Glaucón buscando que lo engañaran? Y quizá fue efectivamente engañado. Pero luego comprobaremos que la idea de que la moral refleja sólo la debilidad y la incapacidad humanas, subvierte su función y hace posible que la sociedad constituya "una empresa cooperativa destinada a la ventaja mutua". Afortunadamente, comprobaremos también que la moral está representada erróneamente en la versión de Glaucón y también en la aparente confirmación de los argumentos de Hobbes y de Hume. En el capítulo final de nuestra indagación vincularemos positivamente una moral basada en el acuerdo con nuestra concepción de los seres humanos ideales como individuos liberales.

1.2 Glaucón describe la justicia como una imposición convencional y social a la naturaleza. Habla de que las personas cometen y sufren injusticas antes de aceptar las restricciones que constituyen la justicia, pero sería más correcto que dijera que, aunque las personas naturalmente hacen aquello que finalmente será considerado injusto, originalmente los individuos permanecen fuera de toda restricción moral, en una situación en la que, como dice Hobbes, "Los conceptos de...justicia e injusticia no tienen cabida".[10] Si la moral es materia de acuerdo, antes de tal acuerdo nada es ni justo ni injusto. Para

Glaucón, el hombre natural es presocial y premoral. Y desde el punto de vista de ese ser presocial y premoral la moral parece un mal necesario.

Alasdair MacIntyre afirma que la concepción de hombre natural "sufre de una fatal incoherencia interna".[11] Los términos utilizados para caracterizar a esa criatura supuestamente presocial, tales como "egoísta" y "agresiva", "se definen atendiendo a normas establecidas de conducta", y por lo tanto constituyen "un vocabulario tomado de la vida social". MacIntyre supone que Glaucón, y la tradición sofista en nombre de la cual habla Glaucón, confunden la sociedad homérica con la naturaleza y al héroe homérico, el *agathos*, con el hombre natural. Glaucón no evalúa la moral desde el punto de vista de la naturaleza, sino que antes bien evalúa una moral, la de la *polis* del siglo V, desde el punto de vista de otro, del *agathos*.

Lo que nos interesa no es la exactitud histórica del ataque de MacIntyre. Para él, como para nosotros, Glaucón desempeña un papel ilustrativo y representativo. MacIntyre sostiene que la filosofía moral occidental ha repetido con frecuencia el error sofista de suponer que existe un punto de vista extramoral, natural o racional, desde el cual puede entenderse la moral y, o bien defenderla o bien atacarla.[12] Debemos negar esta suposición. Pero nuestra negación podría impulsar a MacIntyre a presentar un segundo ataque: que tratar de utilizar en la teoría moral al hombre natural como un punto arquimédico compromete el objetivo de la empresa moral.

A fin de mostrar que la idea del hombre natural es, por lo menos en principio, coherente, deberíamos comenzar por observar que al caracterizar al individuo como un ser asocial, nos preocupamos no por sus orígenes, sino por sus motivaciones y sus valores. Los seres humanos son animales que emplean el lenguaje y se ha dicho que el lenguaje es necesariamente un producto social, al menos en la medida en que su uso exige la internalización de una serie de convenciones. Sin pretender entrar en las complejas cuestiones que plantea esta afirmación, admitamos simplemente que los seres humanos están necesariamente "convencionalizados". Lo que debemos preguntarnos es pues si las convenciones que forman parte de lo que una persona es deben introducir algún elemento de sociabilidad

tanto en sus evaluaciones como en sus motivaciones. Y parece claro que no necesariamente debe ser así.

La condición para hacer una valoración asocial es simplemente no-tuista; el no-tuista no se interesa en absoluto por los intereses de aquellos con los cuales entra en interacción. Su función de utilidad, que constituye la medida de sus preferencias, es estrictamente independiente de las funciones de utilidad de aquellos a quienes él afecta. Pero el no-tuismo no asegura una motivación asocial. Si bien un individuo puede no interesarse en los intereses de los demás, puede interesarse en las situaciones que sólo es posible especificar con referencia a los demás. Michel Laver define sutilmente la condición para la motivación asocial; permitimos "sólo aquellos supuestos relativos a las motivaciones de los individuos que se pueden definir precisamente con referencia a un individuo único".[13] Por consiguiente, al explicar la motivación de la persona x, sólo podemos referirnos a aquellos aspectos de las situaciones que tienen que ver con x y con nadie más.

El carácter asocial no exige que toda motivación satisfaga esta condición estricta. Podemos aceptar otros motivos secundarios socialmente definidos. En un mundo en el que los recursos son escasos, lo cual hace que la demanda agregada de muchos bienes exceda su oferta total, la persona asocial puede sentirse impulsada a adoptar una posición instrumentalmente egoísta y agresiva en relación con los demás. Puede intentar dominar a los otros, como la mejor forma de asegurarse la oferta de los bienes caracterizados asocialmente que satisfagan sus necesidades. Pero su agresividad sólo puede entenderse, lo mismo que cualquier otra disposición social, como un mero instrumento para satisfacer sus intereses básicos, no-tuistas y asociales.

De modo que MacIntyre puede tener razón al afirmar que el hombre natural debe caracterizarse con un vocabulario extraído de la vida social. En la medida en que el hombre natural está en interacción con sus semejantes, ha de manifestar ciertas disposiciones sociales. Pero si esas disposiciones se basan estrictamente en los motivos secundarios del individuo, tales disposiciones son pues compatibles con su carácter asocial básico. De modo que la idea del hombre natural no sufre de una fatal incoherencia interna.

La oposición a la convención moral que manifiesta el hombre natural no supone un conflicto entre normas morales o sociales diferentes, como interpreta MacIntyre cuando afirma que la apelación sofista al punto de vista del *agathos* socava la moral de la *polis*. Antes bien, la oposición está implícita en la idea de que el carácter social es una cuestión de convención, de modo tal que lo que se revela cuando se eliminan todas las prácticas e instituciones sociales, es un ser asocial en sus motivaciones y no-tuista en sus valores. MacIntyre pretende que "los sofistas ...no supieron distinguir... la diferencia entre el concepto de un hombre que permanece afuera de las convenciones de un determinado orden social y desde allí las cuestiona, y el concepto de un hombre que se mantiene fuera de la vida social como tal".[14] Pero, en realidad, los sofistas captaron, por primera vez en el pensamiento humano, el punto de vista de una persona que se mantiene fuera de la vida social, no en su capacidad, no en el sentido de poder vivir sin la sociedad, sino en sus motivaciones, en el sentido de poder considerar la sociedad como un mero instrumento para alcanzar los objetivos que no necesitan de la vida social para ser formulados. Si MacIntyre está en lo cierto al ver al hombre natural de los sofistas como al héroe homérico, los sofistas encuentran pues, haciendo abstracción de las circunstancias sociales del *agathos,* la base para la oposición radical que presentan a la idea misma de un orden moral.

Pero, aunque MacIntyre niegue la coherencia conceptual de la visión medular de los sofistas, parece comprenderla y captar sus implicaciones con gran agudeza. Porque advierte que el argumento sofista hace más que presentar la moral como un mal necesario, aceptado a desgano por aquellos incapaces tanto de dominar a sus semejantes como de vivir sin ellos. Antes bien, sostiene MacIntyre, los sofistas caracterizan la moral como una cortina de humo, un artificio mediante el cual el inteligente puede liberarse de las restricciones mismas que tal artificio parece imponer.

MacIntyre señala

> El hombre natural caracterizado en la figura de Trasímaco [y debemos recordar que Glaucón recoge el argumento de Trasímaco] tiene dos rasgos principales. Su retrato psicológico es sencillo: está dispuesto a tomar todo lo que desea y lo que desea está estrechamente circunscripto. El poder y el placer constituyen sus intereses exclusivos. Pero, para obtener lo que desea, ese lobo debe vestir los ropajes de la oveja de los valores morales convencionales. Sólo puede llevar adelante su mascarada poniendo al servicio de sus fines privados el vocabulario de la moral convencional. Debe decir... lo que la gente desea oír, de modo tal que ésta le ponga el poder en las manos. Por consiguiente... el hombre natural debe aprender el oficio de moldear a la gente mediante la retórica. Debe llevarlos por las orejas antes de atraparlos por el cuello.[15]

Tratar la moral como si fuera el producto de un acuerdo entre personas asociales parecería representarla como un mal necesario. Mac Intyre pinta un cuadro aun más sombrío. El inteligente puede no sólo desembarazarse de las cadenas de la moral, sino engrillar con ellas a sus semejantes. La moral parece así hacer de todos nosotros cooperadores reacios. En realidad nos divide en zorros y gaviotas, es decir, entre taimados y engañados.

1.3 Hay circunstancias en las cuales a una persona le resulta racional sacar ventaja de otra. En ese caso, la razón y la justicia se divorcian y la justicia pierde toda posibilidad de influir en el individuo no-tuista. En 7,5.2 aceptamos gran parte de esta opinión del Necio de Hobbes, pero limitamos nuestra concesión afirmando que la separación de la justicia y la razón es el resultado de la falta de una igual racionalidad de las personas. Y sostuvimos que esa falta es accidental. Pero quizás esto constituya una forma demasiado fácil de rehuir el argumento del Necio. Quizás la falta de una igual racionalidad es algo ideado por aquellos que vislumbran la perspectiva de conservar sus ventajas injustas detrás de una fachada moral.

En su *Discours sur l'inégalité* Jean-Jacques Rousseau introdujo la idea de que una moral basada en un contrato invita a emplearla como un instrumento de dominación. Partiendo de un estado de naturaleza esencialmente hobbesiano, Rousseau sostiene que:

> Especialmente el rico llega pronto a sentir las desventajas de una guerra perpetua en la cual él carga con todos los costos, porque aunque todos arriesguen su vida, los ricos arriesgan además sus bienes... Al carecer de razones convenientes que lo

justifiquen y de las fuerzas suficientes para defenderse, el rico, espoleado por la necesidad, finalmente concibe el proyecto más profundo que jamás haya entrado en el espíritu humano: volver a su favor las fuerzas mismas que lo atacaban, convertir a sus adversarios en sus defensores... "Unámonos", les dice, "para proteger al débil de la opresión, para restringir al ambicioso y para darle a cada uno la posesión segura de lo que le corresponde; instituyamos reglas de justicia y de paz, que todos estén obligados a acatar, sin excepciones, y que han de compensar en cierto grado los caprichos de la fortuna, al obligar del mismo modo al poderoso y al débil a cumplir con los deberes mutuos." ... Tal fue, o podría muy bien haber sido, el origen de la sociedad y de las leyes, con lo cual se consiguió colocarle más cadenas al débil y conferirle nuevos poderes al rico, y establecer para siempre la ley de la propiedad y la desigualdad, se consiguió convertir una usurpación inteligente en un derecho irrevocable y, para beneficio de unos pocos ambiciosos, someter, desde entonces en adelante, a toda la humanidad al trabajo, a la esclavitud y a la miseria.[16]

El acuerdo instituye una sociedad de desiguales en la cual la riqueza confiere poder y finalmente dominio. Al aceptar las restricciones, el rico adquiere un instrumento de dominación y el pobre queda sujeto a un instrumento de represión. Podría cuestionarse la racionalidad de que el pobre se disponga a acatar tales restricciones. Pero la racionalidad misma llega a ser un privilegio del que ejerce la dominación, quien entonces alienta racionalmente la irracionalidad de sus semejantes.

En el capítulo 7 afirmamos que cuando la justicia y la razón se divorcian, la causa debe buscarse o bien en las falsas visiones del mundo y en las prácticas e instituciones a que dan lugar esas cosmovisiones, o bien en las diferencias tecnológicas accidentales que existen entre las sociedades. Ambas causas provocan una racionalidad desigual que impide que haya justicia. Siguiendo el enfoque de Rousseau, que refleja muchas de las ideas sofistas, debemos preguntarnos si aquellos capaces de dominar a sus semejantes no fomentan y mantienen racionalmente dichas causas. No sólo es posible usar la moral contractual como un instrumento de dominación, sino que la razón apoya además ese uso.

En una interesante crítica de la racionalidad instrumental, Mary Gibson sostiene que aquellos que pueden socavar la

racionanalidad de sus semejantes, racionalmente deberían hacerlo. Y afirma que "en una sociedad competitiva, en general, a cada persona le interesa ser lo más racional posible y que sus competidores sean lo menos racionales que puedan ser (por supuesto, con ciertos límites). A los capitalistas les interesa que... los obreros no reconozcan cuáles son sus propios intereses ni obren efectivamente con el fin de satisfacerlos... De ahí que tengamos un sistema en el cual las personas tienen un interés en que *no* se satisfagan las condiciones óptimas de la racionalidad".[17]

Si se satisficieran las condiciones óptimas de racionalidad, siguiendo pues nuestro primer argumento, sería necesaria la coerción para que un individuo o un grupo mantuviera una ventaja injusta sobre los demás. Los costos a largo plazo de la coerción, tanto aquellos directamente relacionados con el mantenimiento del aparato represivo, como aquellos costos de oportunidad provocados por la pérdida de las ocasiones de cooperar, suelen exceder sus beneficios. Pero, si es posible alentar a algunas personas o a ciertos grupos de personas a que se comporten irracionalmente, que sean incapaces de reconocer sus propios interes o de perseguirlos efectivamente, también es posible mantener un orden menos abiertamente coercitivo del cual es fácil sacar ventaja.

Gibson afirma que una sociedad competitiva es "un sistema en el cual las personas que se encuentran en posición de aplicarlos en gran escala se benefician con una gran variedad de fraudes".[18] No sólo se engaña a las personas respecto de sus propios intereses y de los productos que pueden satisfacer tales intereses, sino que los impostores se engañan también a sí mismos "respecto del hecho de que estám montando tales fraudes". Podemos identificar ese autoengaño con el mantenimiento de una ideología en la cual tanto las víctimas como los beneficiarios ocultan la injusticia de sus prácticas. El resultado de ello es que también " se socava el acceso a una información exacta, transmitida de un modo inteligible y que no permita malas interpretaciones, lo cual es esencial para llevar adelante un plan de vida racional".

En tanto ese ocultamiento ideológico sea efectivo, podemos suponer que la idea de una moral basada en el acuerdo, la idea

de una sociedad que funcione como una empresa cooperativa destinada a la mutua ventaja justa, esconde la realidad de una sociedad basada en el acuerdo falso que describe Rousseau. Aquellos que no se dejan engañar por la fachada y que se encuentran en posiciones favorecidas, pueden considerarse exentos de acatar las restricciones de la moral. A ellos la moral ha de parecerles, no como el precio que todos deben pagar para lograr una cooperación efectiva, sino sólo como una restricción que deben acatar aquellos que carecen de la fuerza o la inteligencia para mejorar sus situaciones. Y esto nos remite nuevamente a la tesis desarrollada por Glaucón: entre iguales, la moral sería un mal necesario. Pero en el mundo real, la moral es un instrumento de dominación que presenta como iguales a los que no lo son y como restricción mutua a aquello que sólo es opresión unilateral. Convencionalmente se reconoce que la justicia es superior a la injusticia y, en realidad, una sociedad de personas injustas debe funcionar peor que una sociedad de personas justas. Pero la persona poderosa e inteligente, al reconocer la superioridad natural de la injusticia que tiene éxito y queda impune, puede utilizar las creencias morales de sus semejantes como un medio para desembarazarse de las restricciones que imponen esas creencias.

Rousseau y Gibson parecen confirmar el punto de vista de los sofistas. Si se representa la moral como una restricción convencional a la forma natural de la conducta humana, como un mal necesario antes que como parte integrante del bien humano, se socava pues cualquier adhesión genuina a la moral. En principio uno puede suponer que no hay por qué cocear contra el aguijón, que para obtener lo que desea, un individuo que se sabe débil y que no es autosuficiente, necesita entrar en interacción con los demás y la interacción ventajosa exige restricción. Pero luego ese mismo individuo puede llegar a creer que no es tan débil como parecía. Y puede manipular a los demás, alentándolos a obrar de modos que resulten favorables para los intereses del individuo en cuestión aunque desfavorables para los de ellos. Y lo que es más importante, ese individuo puede utilizar el lenguaje mismo de la restricción como un instrumento para liberarse de ella. Si el interés de esa persona quedara al descubierto lo condenaría, pero al manifestar veneración por la moral el individuo puede sacar ventaja de los otros simulando cooperar con ellos. Apresarlos antes por la oreja que por cuello parece una expresión demasiado cruda; digamos que los toma por la oreja sólo para hacerles volver la cabeza.

2.1 Como ya hemos dicho (porque realmente lo dije), una moral basada en el acuerdo es una moral del hombre económico.[19] Y el hombre económico es el hombre natural de nuestro tiempo. Quien expresó más sucintamente la motivación de ese hombre fue Samuel Gompers, cuya respuesta a lo que exigía la clase obrera fue "¡Más!"[20] El hombre económico, sea capitalista u obrero, es un apropiador indefinido que trata de someter una porción cada vez mayor del mundo a su poder, ejercido al servicio de sus preferencias.

El hombre económico es propiamente una entidad teórica que está presente en nuestras mejores teorías destinadas a conceptualizar y explicar la interacción humana. Pero lo que nos interesa aquí no es su función explicativa. Porque el hombre económico ya forma parte de la esfera de la conciencia de nosotros mismos, del modo en que nosotros, herederos de una tradición de pensamiento social que maduró en la Europa occidental y en Norteamérica desde el siglo XVII, nos entendemos a nosotros mismos. He sostenido en otra parte que nuestra creencia en el hombre económico o en que los seres humanos son radicalmente contractuales es más profunda que nuestros repudios conscientes; el hombre económico es una caricatura, o una deformación que cada vez ha ido dando más forma a nuestra realidad.[21]

Pero, como el heredero del *agathos* homérico, el hombre económico considera que los órdenes moral y político son restricciones convencionales que refrenan la expresión natural de sus deseos. La defensa racional de estas restricciones revela que son males necesarios y, como sostuvimos en la sección anterior de este capítulo, la respuesta natural a aquello que se considera un mal necesario es hacerlo innecesario. Para el hombre económico, una moral basada en el acuerdo es una impostura.

Si la concepción del hombre económico fuera la última palabra sobre nosotros mismos, la moral tendría un futuro precario. Porque las restricciones morales no ejercerían ninguna influencia en aquellos que las consideraran únicamente

como instrumentos de dominación. Al descubrir que las palabras melosas no les rinden ningún provecho, los hombres y mujeres económicos no tendrían otra opción que asir a sus semejantes por el cuello. Suponiendo que la argumentación que desarrollamos en los primeros capítulos fuera acertada, de modo tal que en determinadas condiciones, relacionadas tanto con nuestras circunstancias como con nuestras disposiciones, la racionalidad de las restricciones morales quedara plenamente demostrada, podría ocurrir sin embargo que las características mismas de la naturaleza humana que forman parte del fundamento de tales restricciones sirvieran también para destruir las condiciones en las cuales sería racional acatarlas. Puede parecer que necesitamos las restricciones morales a causa de nuestro carácter asocial y de nuestro no-tuismo, pero que al ser asociales y no-tuistas, obramos con el fin de crear circunstancias en las cuales la supuesta imparcialidad de esas restricciones sea una impostura. Y si se reconoce tal impostura, las restricciones pierden su efectividad. Podríamos suponer que esta conclusión se asemeja mucho al profundo mensaje de Platón y de Rousseau. Sócrates queda hechizado por Glaucón y no le responde.[22]

¿Hemos desarrollado pues un enfoque de la moral apoyándonos en una concepción de la naturaleza humana que a su vez hace que la moral sea una quimera? Negamos esta suposición. Nuestra teoría de la moral, aunque emplee la idea del hombre económico no se compromete con ella porque no la considera una explicación plena y adecuada de la naturaleza humana. El hombre económico continúa siendo una caricatura, aun cuando es una caricatura útil tanto desde el punto de vista explicativo como desde el punto de vista normativo. Una moral basada en el acuerdo le brindaría al hombre económico una restricción beneficiosa siempre que lograra restringirlo. Precisamente porque los seres humanos reales compartimos algunas características del hombre económico, una moral basada en el acuerdo nos brinda una restricción beneficiosa y precisamente porque no somos hombres y mujeres económicos nos dejamos restringir por ella. La moral no se destruye a sí misma.

Y así lo sostendremos. Pero sería bueno recordar que en este caso un argumento no basta. En la medida en que la idea del hombre económico forme parte de la manera en que nos

entendemos a nosotros mismos, forme parte de la idea de qué es ser una criatura humana, en la medida en que esa idea se mantenga firme aun ante el rechazo consciente, las ataduras racionales de una moral basada en el acuerdo carecerán de la fuerza necesaria para contenernos. Además del argumento necesitamos un exorcismo. Pero yo no tengo el poder para realizarlo. Trazaremos los rasgos distintivos del hombre económico y mostraremos cuáles son las virtudes morales y sociales de las que carece. Luego, en el capítulo final, incorporaremos tales virtudes a la idea del *individuo liberal*. Sostendremos que el individuo liberal está dispuesto a aceptar la justicia esencial, que para él la moral no es un mal necesario y mucho menos un instrumento de dominación. Esbozaremos brevemente la idea de una sociedad esencialmente justa que promueve la formación de una comunidad individualista en la que tienen cabida los afectos. Tanto la idea del individuo liberal como la de una sociedad esencialmente justa son conceptos problemáticos y amenazados que han de darle a nuestra conclusión un tono más sombrío.

2.2 El hombre económico siempre quiere más. Pero "más" es un concepto ambiguo; podemos suponer que el hombre económico quiere obtener o bien más de lo que inicialmente tenía o bien más de lo que tienen sus semejantes. La persona que trata de obtener más que los otros, que desea sobrepasarlos, no tiene una motivación asocial. Si sobrepasar a los otros es uno de sus objetivos primarios e intrínsecos, se trata de un individuo motivado por el *amour propre* rousseauniano.[23] Y, como sostiene Rousseau, semejante persona no puede identificarse con el hombre natural asocial, porque sólo existe en relación con los demás y adquiere el verdadero sentido de sí mismo en esa relación.[24] Si el hombre económico es el auténtico heredero del hombre natural de los sofistas, su interés por sobrepasar a los demás sólo puede ser derivado. Al querer más, quiere tener más de lo que tiene en ese momento.

El hombre económico es un apropiador. La relación fundamental que mantiene con las cosas es la expresada en el Génesis, "someterlas y dominarlas".[25] Aunque la apropiación individual excluya a los otros, no es esencialmente una relación social; un individuo se apropia más literal y más completamente de

aquello que luego hace formar parte de su propio cuerpo, de lo que toma para comer y beber. Luego, se apropia de aquello que le sirve de instrumento, de aquello que le permita satisfacer sus preferencias. Las demás personas entran en esta apropiación sólo incidentalmente; en presencia de los otros, el individuo sólo puede asegurarse la apropiación si aquellos limitan el uso que hagan de la cosa apropiada a lo que decida el individuo apropiador.

La apropiación no tiene un límite, un "techo" natural. El hombre económico quiere obtener más. Pero la apropiación, como otras acividades, manifiesta el principio de utilidad marginal decreciente; cuanto más obtiene un individuo, menor es el beneficio y mayor el costo de cada nuevo acto de apropiación —aunque no siempre es así, es lo que ocurre en general—. Un análisis de la relación costo/beneficio determina la extensión efectiva de las apropiaciones del hombre económico.

No debemos suponer que el hombre económico tiene escrúpulos de tipo metafísico o moral contra apropiarse de otras personas. Pero es posible que los otros no sólo se resistan a su actividad de apropiación, sino que además traten a su vez de apropiarse de él. La apropiación expresa una instrumentalidad unilateral; y cuando se vuelve impracticable, el hábitat preferido del hombre económico es el mercado perfectamente competitivo. Porque en el mercado se desarrollan las interacciones emprendidas voluntariamente por personas carentes de toda motivación tuista. Las interacciones del mercado forman una red de arreglos contractuales y esa red se funda en un contrato social general que expresa la instrumentalidad bilateral, la cual, desde el punto de vista del hombre económico, constituye la sociedad. Los "competidores" del mercado, debemos recordarlo, no son ni enemigos ni amigos; no sienten un interés natural mutuo, sino sólo el interés instrumental, artificial, expresado en el contrato.

La idea de caracterizar el mercado con el adjetivo "competitivo" quizás haya sido desafortunada. Porque la palabra "competencia" sugiere una situación de suma cero en la cual el hecho de que una persona gane implica que otra pierda. Pero, como observamos en el capítulo 6, en el mercado perfectamente competitivo cada persona puede esperar mejorar su condición.

La idea sobre la que se basa el mercado no es la de un juego, en el sentido de partida o competencia. El mercado es la estructura dentro de la cual se desarrolla una interacción asocial antes que opuesta; es un medio óptimo para aquellos que deben participar de una interacción mutuamente instrumental. Y el hombre económico se distingue de sus predecesores por cobrar conciencia de ese medio; como veremos más adelante, la sociedad le ofrece más que la mera protección contra sus semejantes.

Al relacionarse con otras personas dentro del contexto del mercado, el hombre económico manifiesta tener una visión *radicalmente contractual* de las relaciones humanas. El hombre económico es un apropiador de las cosas que puede dominar, que puede tratar instrumentalmente sin correr el riesgo de que a su vez otro lo trate instrumentalmente a él; pero es un defensor del contrato cuando tiene que vérselas con personas a las cuales no puede dominar, a las que sólo puede tratar instrumentalmente, si él a su vez les permite tratarlo del mismo modo.

El hombre económico es un defensor del contrato en el sentido en que todas sus relaciones interpersonales libres o no coercitivas son contractuales. Para él, las relaciones sociales voluntarias necesitan tener una base racional; el contrato se las brinda. La idea fundamental del contrato, la idea de que la persona que no se interesa por los intereses de los demás, puede sin embargo entrar en interacción con otras de un modo mutuamente beneficioso, y por consiguiente aceptable, es, como ya observamos, una de las ideas más liberadoras de la historia humana que desembarazó a las personas de la exigencia de depender afectivamente de sus semejantes. Pero el hombre económico lleva esa liberación al extremo más absoluto; su motivación exclusivamente asocial excluye las relaciones interpersonales voluntarias no contractuales. El hombre económico no sólo se liberó de la dependencia afectiva compulsiva; es incapaz de sentir afecto voluntario.

Hobbes ofrece el retrato psicológico del hombre económico más unificado y preciso.[26] Pero la sociedad que genera ese hombre económico con incomparable consistencia partiendo de las motivaciones puramente asociales es primariamente un instrumento destinado a disminuir los costos de la interacción humana, a eliminar la hostilidad que caracteriza la condición

natural de la humanidad. Es cierto que, cuando Hobbes enumera los males de esa condición natural, se refiere a la falta de artes, de erudición y de una sociedad (entendida como el intercambio social), a la carencia de viviendas espaciosas, de naves y de artefactos mecánicos, pero pone el acento principalmente en la corta duración y en la sordidez de la vida. Hobbes no tiene un gran sentido de los beneficios positivos de la interacción, y ningún sentido en absoluto (o por lo menos eso es lo que parece) de los beneficios progresivos que tiene la perspectiva, ofrecida por una condición de sociedad, de un destino para los seres humanos, no meramente mejor, sino en continuo progreso.

La comprensión psicológica del carácter asocial implícita en el argumento de Hobbes no puede por lo tanto equipararse con un grado comparable de visión social o económica. Y los pensadores que siguieron luego la senda de Adam Smith, quien comprendió la idea del progreso social, no lograron igualar la consistencia con que Hobbes caracterizó al hombre económico. El resultado que obtuvieron fue oscurecer las implicaciones que tiene la motivación asocial en la estabilidad de las instituciones y las prácticas sociales.

Cuando Rousseau escribe sobre el contrato que instaura una sociedad de desiguales, ignora, como Hobbes, los beneficios positivos y progresivos que puede proporcionar hasta una sociedad desigual. Por lo tanto supone que la situación de los pobres en realidad empeora como consecuencia de las nuevas relaciones de poder que instituye esa sociedad. Pero podría objetarse plausiblemente que, aunque la estabilidad que proporcionan las nuevas reglas de "justicia y paz" limita ciertas oportunidades de crecimiento individual que antes eran accesibles, el efecto neto que producen en las expectativas es positivo. Cada persona puede razonablemente esperar mejorar su situación como consecuencia de los beneficios del mercado y de la interacción cooperativa, aun cuando esos beneficios no se distribuyan justa o equitativamente. Si entre los hombres económicos las restricciones morales y políticas se convierten en instrumentos de dominación, de modo tal que perpetúan y fortalecen las desigualdades, también pueden servir sin embargo como instrumentos de progreso.

Por consiguiente, hay una importante diferencia entre el hombre económico y sus predecesores. El hombre natural del pensamiento sofista vivía en un medio al que consideraba esencialmente incapaz de progresar. La sociedad misma podía poner orden en los asuntos humanos y el orden en sí mismo era un gran bien. Pero una vez obtenido ese bien y mientras no se socavara la estabilidad de la sociedad, la interacción social constituía un juego de suma cero. Sólo se obtenían ganacias a expensas de las pérdidas de algún otro. Pero en el mundo del hombre económico, el orden es sólo el primer bien proporcionado por la sociedad. Aunque los bienes sean escasos, si existen las instituciones y prácticas sociales adecuadas, la interacción humana puede aumentar indefinidamente la oferta de bienes. Todas las personas pueden entonces mejorar su situación; puede haber ganacias que no impliquen pérdidas.

En la medida en que las restricciones morales contribuyen a establecer el orden sobre el cual se basa el progreso, la moral debe aparecer a una luz más favorable de la que indicaría el enfoque de Glaucón. Glaucón sugiere que las personas aceptan las restricciones morales después de comparar los beneficios de cometer injusticias y los costos de sufrir la injusticia. Pero lo que está en cuestión aquí no son tanto los beneficios y los costos directos de nuestras acciones. Antes bien, aceptamos las restricciones morales a causa de los beneficios indirectos que nos brindan: no el alivio de no tener que sufrir injusticias, sino las ganancias que proporciona la interacción dentro de una estructura ordenada y estable. Las restricciones a la fuerza y el fraude hacen posible el mercado, pero nadie sufre injusticia, sencillamente porque la injusticia está ausente.

Con todo, aun habiendo demostrado que la moral es más que un mal necesario, parecería que no la hemos liberado de ser un instrumento de dominación. Todo lo que muestra nuestro argumento es que lo que está detrás de la fachada es algo más que una simple explotación unilateral; la moral ha demostrado ser más que un artificio destinado a permitirle a un grupo beneficiarse a expensas de otro. Aquellos que pueden utilizar el lenguaje de la restricción como un instrumento para desembarazarse de tal restricción, no han de negarles por ello a sus semejantes los beneficios del orden y el progreso; sencillamente

se aseguran de recibir la porción más abultada de tales beneficios.

2.3 Puede suponerse que, al tratar de dominar y explotar a sus semejantes, el hombre económico intenta socavar la racionalidad de los otros. Y eso es en realidad, lo que sostiene Mary Gibson cuando dice que "en una sociedad competitiva... cada persona está interesada en que sus competidores tengan la menor racionalidad posible".[27] Pero semejante afirmación es demasido fuerte, pues le concede una importancia excesiva a la visión de suma cero de la competencia humana, lo cual puede haberles parecido evidente a los héroes homéricos, pero que los hombres económicos que aprecian el mercado aprendieron a rechazar. El mercado, que sólo florece en ausencia de la fuerza y el fraude, permite desarrollar a través del tiempo un movimiento en dirección deseable, y alienta no sólo el acomodamiento estático de determinadas preferencias, sino también un acomodamiento dinámico y flexible a medida que van apareciendo nuevas preferencias y nuevas tecnologías de satisfacción. Pero es necesario contar con un alto grado de racionalidad instrumental para entender la disciplina por encima de la gratificación inmediata, que es una condición previa de ese progreso. Seguramente podríamos imaginar seres que se preocupen poco o no se preocupen en absoluto por su futuro y que por lo tanto podrían no interesarse en esa disciplina; pero si suponemos que al hombre económico lo motiva por lo menos un grado moderado de prudencia, o un interés por sí mismo a largo plazo, debemos suponer también que ese hombre ha de encontrar esa prudencia en su interés por participar en una interacción con personas cuya racionalidad les permita comprender cuáles son los medios para satisfacer tal prudencia. La irracionalidad de los demás amenaza con echar por tierra los planes y esfuerzos del individuo.

Por lo tanto, la crítica que hace Gibson está errada. Pero acierta cuando dice que el hombre económico tiene un interés meramente instrumental por la racionalidad de los demás. Y esto lo lleva a tratar de socavar la capacidad plenamente reflexiva y crítica de los otros, caracterizada en el capítulo 2. Si la coerción abierta es un instrumento demasiado crudo e ineficiente y fracasan las palabras almibaradas, el hombre

económico influye en los demás manipulando sus preferencias y, al hacerlo, llega a adoptar una visión de la justicia que convierte la moral en un instrumento de dominación, en un sentido más profundo del que hemos considerado hasta ahora.

Al extender el carácter asocial de sus predecesores al hombre económico, debemos cuidarnos una vez más de no negar la función que cumple la sociedad al darle forma a los valores y las motivaciones asociales y no-tuisticas del hombre económico. Y no debemos suponer que el hombre económico ignora los efectos de la socialización. Le importa la sociedad como un instrumento o una herramienta que le permite alcanzar sus fines personales, una herramienta que, le guste o no le guste, debe compartir con los otros. Pero, el hombre económico es consciente de que, al permitirle a las personas alcanzar más fácilmente sus fines asociales, las instituciones y prácticas sociales también reaccionan y dan nueva forma a esos fines. Al concebir a las personas como conjuntos de preferencias, representados por funciones de utilidad, y conjuntos de capacidades y posesiones, representados por funciones de producción, el hombre económico advierte que ni las funciones de utilidad ni las funciones de producción son algo fijo, advierte que el uso del instrumento mismo que les sirve puede alterarlas.

Una sociedad justa maximiza el beneficio relativo mínimo entre sus miembros. Si consideramos las preferencias y las dotes de los individuos como algo dado, podemos pues utilizar los principios de justicia para determinar y evaluar las instituciones y prácticas sociales. Pero podemos obrar en la otra dirección. No podemos sencillamente considerar las instituciones y prácticas sociales como algo dado y utilizar los principios de justicia para determinar las preferencias y las dotes individuales, puesto que debe considerarse que las instituciones y prácticas sociales determinan directamente las preferencias y las dotes, y esa determinación puedo satisfacer o no satisfacer las exigencias de la justicia. Antes bien podemos preguntarnos qué instituciones y prácticas sociales satisfacen los requisitos de la justicia, no maximizando directamente el beneficio relativo mínimo de determinados individuos, sino estableciendo las preferencias y dotes individuales de modo tal que se logre el beneficio relativo maximínimo.

En *Un mundo feliz* de Aldous Huxley tenemos el ejemplo de una sociedad diseñada para modelar a los individuos de modo tal que éstos se adapten a sus instituciones y prácticas y cumplan las funciones que exige la supervivencia de tal sociedad.[28] Huxley, como muchos novelistas antiutopía, presenta arbitrariamente una falla en los procedimientos de socialización, pero por el momento podemos suponer que lo hace con éxito. La sociedad demuestra ser el instrumento perfecto de sus miembros, pues los diseña precisamente para que ellos la consideren el instrumento perfecto. En la *República* ,Sócrates sostiene que en la ciudad justa, cada uno hace aquello a lo que mejor se adapta naturalmente.[29] El "valiente nuevo mundo" da vuelta por completo esta concepción; es justo precisamente porque cada individuo está hecho para adaptarse a lo que hace. No necesariamente tenemos que suponer aquí que hay una sola sociedad que cumple con esta norma. Lo que todo "valiente nuevo mundo" exige para ser justo es que se satisfaga una condición de coherencia: las preferencias y capacidades creadas por las instituciones y prácticas sociales deben ser idénticas a las preferencias y capacidades que tales instituciones y prácticas mejor satisfagan. Digamos que una sociedad que diseña sus procedimientos de socialización a fin de que satisfagan esa condición de coherencia es *artificialmente justa.*

La idea de una sociedad artificialmente justa puede parecer muy alejada de la visión puramente instrumental de la sociedad sostenida por el hombre económico. La justicia artificial parece ser la perfección del colectivismo, de la idea de dar forma al individuo para que se ajuste a la sociedad en su conjunto, mientras que el hombre económico individualista puro a quien sólo le interesa la sociedad en la medida en que ésta esté concebida para satisfacerlo. Sin embargo, las dos concepciones están estrechamente relacionadas entre sí. El ideal de instrumentalismo social es simplemente la adaptación perfecta entre las preferencias individuales y las prácticas sociales. Pero nada especifica cómo debe lograrse esa adaptación perfecta. El hombre económico no le da otro valor a las preferencias que el de material destinado a su satisfacción. Por lo tanto no tiene un interés intrínseco en manipular las preferencias o a las personas con el fin de que éstas se adapten a las prácticas sociales. En

realidad, sólo desde el punto de vista del manipulador la manipulación es relevante para el hombre económico. Y aquí podemos advertir la evidente apelación a la justicia artificial.

Concebimos el mercado como un productor de bienes destinados a satisfacer a las personas. Pero el anunciante puede pensar en producir personas a las cuales satisfagan esos bienes. El hombre económico se manifiesta en las actividades de semejante anunciante manipulador casi tanto como en las actividades del productor de bienes de consumo. ¿Qué diferencia hay entre crear una demanda para una oferta dada y crear una oferta para una demanda dada?

Para nosotros, la diferencia estriba en que en una sociedad artificialmente justa los individuos deben carecer de la capacidad de cambiar de ciertos modos humanamente posibles. Deben tener preferencias fijas en dos sentidos: como rendimientos de la razón social y como contribuciones destinadas a la satisfacción social. Debe eliminarse cualquier variación que pueda destruir la adaptación de ambos. Pero entonces podemos preguntarnos si esos habitantes del "valiente nuevo mundo" son plenamente racionales. Actúan racionalmente con el fin de maximizar sus utilidades. Pero, las preferencias cuya medida tratan de maximizar, ¿son plenamente reflexivas? ¿Poseen esas personas una completa capacidad de autocrítica, que les permita someter sus preferencias a una evaluación reflexiva y alterarlas, y por lo tanto someter su conducta basada en las preferencias a la luz de esa evaluación?

Si cada individuo puede someter sus preferencias a una reflexión autocrítica sin afectar la satisfacción justa y óptima que recibe de una estructura social dada, "el valiente nuevo mundo" debe pues formar a sus miembros de modo tal que cada uno apoye reflexivamente el *statu quo* de esa sociedad. No necesariamente debemos considerar si en principio esto es posible. Porque estamos considerando el diseño humano de una sociedad y el diseño social de los seres humanos dentro de los límites de la experiencia posible desde el punto de vista tecnológico y desde el punto de vista imaginativo. Y podemos llegar a la conclusión de que sólo puede asegurarse la adaptación entre las preferencias producidas socialmente y aquéllas satisfechas socialmente suprimiendo la racionalidad reflexiva plena. El

defecto de *Un mundo feliz* de Huxley, que ahora podemos señalar de manera no arbitraria, es que no se ha eliminado suficientemente esa capacidad de reflexión.

En la medida en que una sociedad da forma al individuo, considerar que la sociedad es una mera herramienta induce a considerarla una herramienta destinada a dar forma a los individuos. Y si se supone que es una herramienta su eficacia se mide desde el punto de vista del usuario. Al juzgar atendiendo a sus intereses asociales, el hombre económico trata de utilizar la herramienta, no para producir seres humanos con una racionalidad reflexiva, sino antes bien para producir criaturas instrumentalmente racionales que utilizan la herramienta para satisfacer sus preferencias. La justicia que persigue el hombre económico, la adaptación de las personas a la sociedad, es artificial y al procurarla, podemos legítimamente suponer que amenaza la racionalidad plena de sus semejantes.

2.4 Aunque el hombre económico encuentre en la sociedad una fuente de beneficios continuados y crecientes no reconocida por sus predecesores anteriores a la aparición del mercado, su motivación asocial restringe severamente el alcance de tal beneficio. Porque un actor asocial procura sólo objetivos que puedan especificarse desde el punto de vista de una sóla persona: él mismo. Por consiguiente no puede obtener placer o satisfacción directos de las actividades esencialmente sociales. Obsérvese que para conseguir placer de una actividad no es necesario tener motivaciones tuísticas. Puedo obtener placer conversando con un amigo. Esta es una actividad esencialmente social. Puedo obtener placer aun cuando la otra persona no sienta lo mismo. Pero aun en el caso de que yo fuera indiferente respecto del placer que siente la otra persona, mi motivación no es asocial. No puede especificarse la fuente de mi placer sin referencia a la otra persona; a una persona diferente de mí.

Hay muchísimas actividades de alcance social aun cuando la satisfacción que brindan sea puramente individual e independiente. En muchos y diversos casos, desde las actividades atléticas a las musicales, desde las actividades políticas a las militares, los seres humanos sienten satisfacción por el simple hecho de participar. Seguramente la satisfacción debe estar relacionada con algún fin, pero sería errado concebir la participación como un valor puramente instrumental. En realidad, un individuo puede valorar los fines en parte porque éstos le proporcionan ocasiones de participar de un esfuerzo conjunto o cooperativo. Y la persona que aprecia la participación no considera que su falta de autosuficiencia sea una debilidad. En realidad, considera que una vida autosuficiente no es plenamente humana.

Razonando desde un punto de vista evolutivo, podemos suponer que el valor que le damos a la participación está relacionado instrumentalmente con nuestra insuficiencia. Podemos coincidir con lo que dice Hume: "Si cada individuo poseyera todas las facultades necesarias tanto para su propia preservación como para la propagación de su especie", seríamos incapaces de desarrollar cualquier forma de actividad social.[30] Pero esto no hace que el valor de la participación sea en sí mismo instrumental. Una explicación instrumental de algunos de nuestros valores intrínsecos no tiene nada de contradictorio, ni de paradójico; ni siquiera nada sorpendente. Al advertir que no somos autosuficientes valoramos la participación con nuestros semejantes, pero la valoramos por lo que es y no como un medio para superar esa insuficiencia.

Al comprender esto, deberíamos advertir hasta qué punto el hombre económico es ajeno al yo real. Porque su motivación asocial lo incapacita para darle a la participación un valor más que instrumental, o de manera más general, no sólo a la participación sino también a cualquier actividad que exija la interacción con sus semejantes. De modo que su falta de autosuficiencia es sencillamente un mal que hace que para lograr su propia satisfacción dependa de otras personas, con intereses diferentes de los suyos. Para él, la participación, lejos de constituir un bien, llega a ser simplemente un recordatorio de su dependencia. El hombre económico apoya el enfoque de Rousseau en lo referente a que la división del trabajo disminuye la libertad del individuo, porque lo obliga a depender de la voluntad de los demás. [31] La idea de que la división del trabajo puede ser una fuente tanto de enriquecimiento como de eficiencia, es decir que puede crear nuevas fuentes de satisfacción, claramente supone la posibilidad de una motivación social. Volveremos a referirnos a este tema cuando examinemos la

comunidad individualista con una sociedad intrínsecamente justa.

Al no darle valor a la participación, el hombre económico también considera que la moral tiene un valor instrumental disminuido. Porque la moral, como la entendemos, suministra una serie de restricciones peculiarmente conducentes a valorar la actividad participativa. Una moral basada en el acuerdo garantiza una distribución racionalmente aceptable de los bienes logrados en tal actividad. Y por lo tanto, el participante no se considera victimizado. No tiene necesidad de sacrificar la porción justa que le corresponde de otros bienes a fin de obtener las satisfacciones que le brinda la participación. Y no obtiene más que una porción justa, de modo tal que no tiene por qué temer el resentimiento justificado de sus semejantes.

En el próximo capítulo analizaremos más profundamente estas ideas. Lo que nos importa aquí es destacar lo que le falta al hombre económico y relacionar lo que le falta con lo que le proporciona la moral. En realidad, una moral basada en el acuerdo quizá sea la única moral que pueda comprender el hombre económico, pero su indiferencia le hace disminuir el valor que tiene la moral en muchas de las actividades que permite realizar.

2.5 La moral no compromete los afectos del hombre económico. Al no interesarse por los intereses de los demás, tampoco se interesa por las restricciones que derivan de tales intereses, excepto en la medida en que la adhesión a tales restricciones le sirva de instrumento para sus propios intereses asociales. Ahora bien, puede parecer que ese desinterés afectivo por la moral es un supuesto de nuestra teoría o de cualquier otra teoría contractual que considera que la racionalidad de los principios morales es estrictamente no-tuista. Pero esto sería un error. Así como los seres humanos pueden considerar intrínsecamente valiosa la participación con sus semejantes, aun cuando sus actividades participativas tengan su base racional y su explicación en la debilidad y en la falta de autosuficiencia, del mismo modo los seres humanos pueden sentirse comprometidos afectivamente con la moral misma, aun cuando sus principios morales tengan una base racional no-tuista.

Debemos distinguir dos formas diferentes en las cuales

pueden concebirse las consideraciones morales que han de comprometer nuestros afectos. Primero, podríamos suponer que los seres humanos son capaces de sentir una moral afectiva. Tal moral puede concebirse como una serie de restricciones no a la maximizaciones de la utilidad del actor, sino antes bien a la maximización de sus intereses asociales, motivados por el interés que tiene por sus semejantes. Las restricciones no comprometen racionalmente en virtud de cualquier interés o cualquier preferencia del individuo, sino que son vinculantes en virtud y sólo en virtud de ciertos intereses directamente tuistas. El hombre económico carece de tales intereses y nosotros hemos evitado cualquier apelación a tales intereses al desarrollar una base racional de la moral. No negamos que las personas reales, a diferencia de los hombres económicos, están limitadas por intereses tuistas en la búsqueda de sus intereses asociales y, por supuesto, no negamos que pensadores de la moral tales como David Hume desarrollaron explicaciones de la moral afectiva basada en tales intereses. Pero una moral afectiva, una moral de los sentimientos, de la benevolencia, nos compromete de un modo muy diferente del de una moral racional; de una manera tan diferente que no hay conexión perceptible entre ambas. Nos gustaría que pudiera reformarse el vocabulario que vincula esas dos formas de compromiso en nuestro esquema conceptual.

Pero no por ello debemos rechazar toda conexión entre la moral y los afectos. Porque no tenemos que confundir la capacidad de acatar una moral llamada afectiva con una segunda forma en que nos comprometen las consideraciones morales, mediante una capacidad afectiva de acatar la moral (racional). El hombre económico carece de esta capacidad, como también carece de la capacidad de valorar la participación. Carece de cualquier interés afectivo por las restricciones aunque sin embargo, puede reconocer que son racionalmente aceptables. Una moral basada en el acuerdo sólo puede hablarle a su intelecto, no a sus sentimientos.

Las personas que tienen una capacidad afectiva para aceptar la moral sienten que sus emociones y sus sentimientos están comprometidos con aquellos que reconocen como consideraciones morales. La capacidad misma puede ser neutral respecto de la naturaleza o la fuerza vinculante de esas considera-

ciones; aunque lo que nos interesa son las restricciones basadas en la racionalidad, podemos suponer que aquellos que creen en una autoridad objetiva pueden sentir que sus emociones están comprometidas con las demandas que le haga tal autoridad. Una persona que posee sentido del deber manifiesta una capacidad afectiva por acatar la moral; se siente movida a cumplir con su deber sólo porque es su deber. Una persona que posee sentido de la justicia también manifiesta una capacidad afectiva para acatar la moral; se siente movida a hacer lo que es justo sólo porque considera que es justo. Obsérvese que una capacidad afectiva para aceptar la moral supone una concepción previa de la moral; una persona no puede sentirse movida por el sentido del deber, salvo que anteriormente crea que cierta acción constituye su deber.

Al distinguir entre una capacidad afectiva para aceptar la moral y la capacidad para acatar una moral llamada afectiva, es necesario aclarar un poco más este último punto. La capacidad para aceptar una moral afectiva no supone ninguna concepción previa de la moral. Es sencillamente la capacidad de restringir la propia búsqueda asocial a causa de un interés por los otros. Puede decirse que es una moral por cuanto presenta restricciones, pero no porque induzca al individuo a acatar restricciones que ya reconoce. Por el otro lado, una capacidad afectiva para aceptar la moral no presenta restricciones, pero dispone al individuo a acatar, por sus emociones y por sus motivaciones, las restricciones aceptadas previa e independientemente.

Glaucón afirma que si el hombre justo se colocara el anillo de Giges no se conduciría de manera diferente que el hombre injusto.[32] Al decir esto Glaucón concibe meramente al hombre justo como alguien que reconoce la necesidad de aceptar ciertas restricciones, pero cuyas emociones no se sienten en absoluto comprometidas con tales restricciones. Por consiguiente, el "hombre justo" carece de todo sentido de justicia, de toda capacidad para sentirse movido por las consideraciones de justicia como tales. No resulta sorprendente pues que, teniendo el anillo de Giges, se comportara como lo hizo el pastor de Lidia. Pero esto lo único que muestra es que en realidad éste no era el hombre justo. Cabalmente entendido, el hombre justo es la

persona que, reconociendo que cierto curso de acción es justo, se siente emocionalmente comprometida por ese reconocimiento y por lo tanto se siente movida a adherirse a ese curso de acción porque es justo. Seguramente, esa persona podría sentirse movida de otros modos, pero para la persona justa esa motivación es esencial.

El hombre económico carece de la capacidad de ser auténticamente el hombre justo. Comprende los argumentos de la restricción moral, pero la considera un mal del que podría liberarse. En cuanto tiene la oportunidad de utilizar la moral como un instrumento de dominación, sin dudarlo un instante lo hace, porque el interés que tiene por la moral es puramente instrumental y sus objetivos (a los cuales la moral sirve de instrumento) son asociales.

Una moral basada en el acuerdo tiene un fundamento racional no-tuista. Sus restricciones obligan desde el punto de vista racional e independientemente de toda preferencia particular. Pero esto no equivale a decir que una moral basada en el acuerdo obligue solamente a los no-tuistas. Aquí podemos hacer la distinción entre el no-tuista puro, cuyos afectos sólo se sienten comprometidos con sus propios intereses, y la persona que exige que la restricción tenga una base racional no-tuista, pero que cuando se cumple esa condición puede sentirse comprometida con tales restricciones. Este último no se permite abrigar sentimientos tuistas que lo pongan en situación de desventaja o que empeoren su posición, pero tampoco trata de sacar ventaja de los demás o de mejorar su situación respecto de ellos evadiendo las restricciones que no estén justificadas desde el punto de vista no-tuista.

A lo largo de todo nuestra exposición hemos empleado el no-tuismo como un supuesto necesario para asegurar que la moral no depende de los afectos, es decir que el no-tuismo le habla directamente a la razón y no a sentimientos o emociones particulares, contingentes. No hemos dicho que el no-tuismo caracterice literalmente a los seres humanos. De modo que podemos suponer ahora que las personas tienen, en particular, una capacidad afectiva para acatar la moral, lo cual no significa que ciertas restricciones morales estén justificadas, sino antes bien que, para algunas personas que poseen esa capacidad las

restricciones morales no constituyen un mal necesario que hay que evitar voluntariamente para convertir la moral en un instrumento de dominación. Ahora estamos en condiciones de refutar la acusación de que la moral basada en el acuerdo es sólo una fachada detrás de la cual los seres no-tuistas o asociales se escudan para perseguir más eficazmente sus fines privados. Ahora estamos en condiciones de mostrar que una moral basada en el acuerdo es algo más que la moral del hombre económico.

Notas

1 Platón, *Republic*, 359c-360b.
2 *Ibid.*, 360b-c; traduc. A. Bloom, pág. 38.
3 *Ibid.*, 358e-359b, págs. 36-37.
4 Hobbes, *Leviathan*, capítulo 15, pág.79.
5 *Ibid.*, cap. 31, pág. 187.
6 Véase mi "David Hume, Contractarian", *Philosophial Review* 88, 1989, págs. 3-38.
7 Hume, *Enquiry*, iii, i, págs. 190-191.
8 *Ibid.*, iii, i, págs. 191-192.
9 Véase el himno al deber de Kant en *Critique of Practical Reason*, págs.86-89.
10 Hobbes, *Leviathan*, capítulo 13, pág. 63.
11 A. MacIntyre, *A Short History of Ethics*, Nueva York, 1966, pág. 18.
12 Véase *ibid.*
13 M. Laver, *The Politics of Private Desires*, Harmondsworth, Middx., 1981, pág. 13.
14 MacIntyre, *op. cit.*, pág. 18.
15 *Ibid.*
16 Jean-Jacques Rousseau, *Discours sur l'inégalité*, 1755, en *The Political Writings*, comp. de C.E. Vaughan, 2 vols. Cambridge, 1915, vol. i, págs. 180-181; mi traducción.
17 Mary Gibson, "Rationality", *Philosophy and Public Affairs* 6, 1977, pág. 218.
18 *Ibid.*
19 "Reason and Maximization", *Canadian Journal of Philosophy* 4, 1975, pág. 433.
20 Véase G. Seldes, *The Great Quotations*, Nueva York, 1966, pág. 285.
21 Véase "The Social Contract as Ideology", *Philosophy and Public Affairs* 6, 1977, especialmente las págs. 135-138.

22 Aquí se oculta una interpretación de la *República*, pero no es la ocasión adecuada para desarrollarla.
23 Véase Rousseau, *Discours sur l'inégalité*, nota o, pág. 217.
24 Véase *ibid.*, pág. 195.
25 Véase Génesis 1: 28.
26 Véase la introducción a *Leviathan de* C. B. Macpherson, Harmondsworth, Middx., 1968, págs. 30-39.
27 Gibson, *op. cit.*, pág. 218.
28 Véase A. Huxley, *Brave New World*, Londres, 1932.
29 Véase *Republic*, 433a-434c.
30 Hume, *Enquiry*, iii, i, pág. 191.
31 Véase Rousseau, *Discours sur l'inégalité*, págs. 169-179. Y también mi análisis en "The Politics of Redemption", *Trent Rousseau Papers*, comp. de J. McAdam, M. Neumann y G. Lafrance, Ottawa, 1980, especialmente las págs. 73-80.
32 Véase *Republic*, 360b-c.

XI

El individuo liberal

1.1 En un pasaje que recuerda curiosamente la descripción que hace Hobbes del estado de naturaleza, Skepticus, uno de los discípulos del Saltamontes de Bernard Suits, dice: "En Utopía el hombre no puede trabajar, no puede administrar ni gobernar, allí no hay arte ni moral ni ciencia ni amor ni amistad".[1] Skepticus sugiere que "Quizás el ideal moral del hombre sea sólo un supremo orgasmo", pero el Saltamontes muestra que semejante propuesta es por completo insostenible. Utopía, la condición en la cual "se han eliminado todas las actividades instrumentales de los seres humanos",[2] parecería algo tan desolado como el estado de naturaleza, en el cual esas actividades instrumentales están condenadas a fracasar.

Al principio, el Saltamontes piensa que no. Porque lo que queda, cuando se eliminan todas las actividades instrumentales son juegos y "jugar un juego es tratar voluntariamente de superar obstáculos innecesarios".[3] Y si los juegos son posibles también lo es el esfuerzo y con el esfuerzo "vuelve a ser posible admirar y compartir, y por lo tanto también el amor y la amistad...[con lo cual] se obtiene la satisfacción emocional del arte. Y quizás hasta la moral esté presente... en la forma de aquello que ahora llamamos espíritu deportivo".[4] Lo que vislumbra el Saltamontes es una cultura basada, no como la nuestra y todas las culturas pasadas, en la escasez, sea ésta "económica, moral, científica o erótica", sino en la plenitud. Esta es una cultura cuyas "instituciones... han de fomentar el deporte y los juegos..., instituciones que han de exigir ... mucha más energía de la que dedicamos hoy a servir a las instituciones de la escasez".

Pero luego el Saltamontes tiene una visión "de la ruina de Utopía, una visión del paraíso perdido. Vi [dice] que con el paso del tiempo los Buscadores y los Trepadores llegaban a la conclusión de que si sus vidas eran meramente juegos, apenas si valía la pena vivirlas".[5] De modo que se engañan y prefieren

creer que sus juegos "son tareas de vital importancia que es indispensable realizar para que la humanidad pueda sobrevivir. Por consiguiente, aunque todas las actividades aparentemente productivas del hombre sean juegos, no se las considera como tales... Y si fuera posible convencer a esas personas de que en realidad sólo han estado jugando, éstas sentirían que sus vidas no han sido nada: una mera obra de teatro o un sueño vacío".

Antes de poder reflexionar sobre el miedo expresado en su visión —el miedo de que el paraíso sólo pueda ganarse para después perderse—, el Saltamontes muere y el libro de Suit termina. Pero ese temor parece tener bases firmes. Al considerar por qué, seguramente obtendremos una comprensión más profunda del bien de la justicia, una comprensión que ha de revelarnos que las restricciones de la moral no son en absoluto una impostura.

Al analizar el mensaje que transmite el Saltamontes, hay dos puntos que merecen especial atención. El primero es que Utopía supone la completa ausencia de la escasez. Y aquí no nos referimos únicamente a los factores materiales; el Saltamontes considera que nuestro mundo actual está desprovisto de emociones. La persona que busca la satisfacción emocional es un buscador tan sumergido en una actividad instrumental como lo es la persona que busca calmar el hambre o la sed. Un aspecto más controvertido es quizá la afirmación del Saltamontes de que la falta de conocimiento es también una forma de escasez; la búsqueda de conocimiento, como la actitud opuesta al goce que produce la contemplación, es también una actividad inadecuada para Utopía.

Ahora bien, al descartar de Utopía todas las formas de actividad instrumental, el Saltamontes sigue una distinguida tradición de la cual Aristóteles es uno de los primeros representantes.[6] Todas aquellas actividades que Aristóteles considera no merecedoras de la perfección son intereses instrumentales que no se adaptan a Utopía. Pero el Saltamontes y Aristóteles toman caminos divergentes cuando consideran qué queda cuando se elimina lo meramente instrumental. Para Aristóteles, la única forma de actividad perfecta es la contemplación. Pero una vida de contemplación, por más apropiada que pueda ser para

el Motor Inmóvil, es —y él mismo lo admite— demasiado elevada para el hombre. Por su parte, el Saltamontes en ningún momento supone que los seres humanos puedan dedicarse en Utopía a una actividad cuyo carácter fundamental difiera de aquellos que conocemos: una actividad relacionada enteramente con la posesión, antes que con la búsqueda o el esfuerzo. Por el contrario, el Saltamontes afirma que la condición utópica de los seres humanos debe consistir en realizar el mismo tipo de actividades que antes realizaban, pero sin la instrumentalidad objetiva que los guiaba anteriormente; la búsqueda y el esfuerzo deben relacionarse con obstáculos innecesarios impuestos voluntariamente. Las actividades utópicas pueden parecer similares a aquellas a las que actualmente se dedican las personas, sólo que en Utopía serían juegos. De modo tal que los investigadores científicos estarían en la posición de estudiantes que se esfuerzan por resolver problemas cuyas respuestas están escritas al final del libro.

Podemos expresar la tesis que subyace en lo que declara el Saltamontes, diciendo que lo que prueba el valor intrínseco de la vida humana es la dedicación a actividades instrumentales: la búsqueda y el esfuerzo. En una condición no utópica, esas actividades también tienen un valor instrumental, puesto que no las perseguimos por puro juego. Sólo que en Utopía ese valor instrumental queda anulado; allí tales actividades tienen el único valor intrínseco de los juegos.

El segundo punto que merece atención es el que surge de la visión del Saltamontes. Porque en ella la tesis sobre el valor se extiende de un modo que resulta fatal para la posibilidad de una verdadera Utopía. Lo que tiene valor intrínseco de la vida humana es no la dedicación a las actividades instrumentales, sino la dedicación a actividades que tienen un valor instrumental. Si no hay otro objetivo más que la búsqueda y el esfuerzo, buscar y esforzarse es "una mera obra de teatro o un sueño vacío". En nuestro mundo no utópico los juegos tienen un objetivo porque son parte de una vida que no es un mero juego. Nos ofrecen una tregua de las actividades relacionadas con los obstáculos necesarios. De modo tal que para nosotros los juegos tienen un valor instrumental; disfrutamos de la libertad de crear nuestros propios obstáculos. Pero en Utopía no tendríamos otra cosa que

esa libertad. En nuestro mundo real el objetivo de los juegos es no tener ningún otro objetivo más allá de ellos; en Utopía los juegos carecerían de ese objetivo.

Pero, ¿no nos encontramos aquí ante una paradoja? ¿Cómo puede el valor intrínseco depender del valor instrumental? ¿Cómo es posible que lo único que vale la pena hacer por simple juego sea hacer algo por algún objetivo adicional? Resulta evidente que la visión que tiene el Saltamontes del valor intrínseco —o por lo menos la interpretación que nosotros hacemos de ella— es completamente incoherente. Para evitar esa acusación debemos complicar levemente nuestro análisis. Ciertamente las actividades que tienen un valor instrumental deben poseer un objetivo más allá de ellas mismas y ese objetivo no puede ser la mera dedicación a tales actividades. Debemos admitir por lo tanto que existen ciertos estados de cosas intrínsecamente valiosos, deseados sencillamente por lo que son. Pero no debemos necesariamente suponer que esos estados de cosas, aunque sean valiosos, bastan para que los seres humanos consideren que la vida merece ser vivida. Podemos aceptar que una vida de búsqueda y esfuerzo —si eso es lo que se considera una vida que vale la pena vivir— exigiría un fin que se pueda alcanzar y poseer. También podemos aceptar que una vida de logros y posesiones —si también eso es lo que se considera una vida que vale la pena vivir— exigiría la búsqueda y el esfuerzo. De modo que podemos decir que la búsqueda y el esfuerzo son intrínsecamente valiosos sólo en la medida en que constituyan objetivos reales, de modo tal que su valor intrínseco dependa de que tengan un valor instrumental, que a su vez exige que otras cosas tengan valor intrínseco.

Las declaraciones que hacemos aquí son claramente psicológicas. Se refieren al bien humano y las cuestiones relativas al bien humano exigen una investigación empírica. Lo que pretendemos no es confirmar estas declaraciones, sólo queremos sugerir que son plausibles. Porque nos permiten mostrar cómo, para aquellos seres que las consideran ciertas, es posible una moral que no sea ni un mal necesario ni un instrumento de dominación. Quizá no seamos esa clase de seres, pero quizás —y yo lo creo— sí.

1.2 Si una vida humana satisfactoria debe incluir activida-

des con valor instrumental, sólo puede alcanzarse el paraíso para después perderlo. Se gana el paraíso cuando se superan todos los obstáculos que se oponen a esa satisfacción; pero cuando se logra salvar todos los obstáculos, las actividades instrumentales pierden su objetivo y dejan de proporcionar satisfacción. Y al perder el paraíso llegamos a tener una nueva comprensión de la función que cumple la escasez en los asuntos humanos. En el sentido más amplio, es la escasez lo que da lugar a las actividades con valor instrumental. Si esas actividades son necesarias para la realización humana, la escasez también es necesaria. La idea de una sociedad basada no en la escasez sino en la plenitud es una quimera; superar la escasez sería superar las condiciones que le dan un objetivo a la vida humana.

Karl Marx le rindió el más elevado homenaje a la sociedad capitalista al suponer que había logrado superar la escasez mediante las fuerzas productivas. Para Marx, el comunismo pudo surgir sólo porque la humanidad pudo pasar de una condición de escasez a una de plenitud. Pero Marx estuvo errado. Con todo, en un pasaje muy criticado de *La Ideología Alemana*, en el cual por única vez Marx especula sobre el carácter de una sociedad de plenitud, el autor revela una visión profunda de lo que sería necesario para crear esa condición.

El efecto que produce la escasez en los asuntos humanos es provocar la división del trabajo, en la cual

> cada hombre tiene una particular, exclusiva esfera de actividad que se le ha impuesto y de la que no puede escapar. Es un cazador, un pescador, un pastor, un crítico y debe continuar siéndolo si no quiere perder sus medios de vida; mientras que en una sociedad comunista, en la que nadie tiene una esfera exclusiva de actividad, sino que cada uno puede realizarse en cualquier terreno que desee, la sociedad regula la producción general y por lo tanto me permite hacer algo hoy y otra cosa mañana, cazar por la mañana, pescar por la tarde, arrear el ganado al atardecer, criticar después de cenar, según las ganas que tenga, sin transformarme en un cazador, un pescador, un pastor o un crítico.[7]

Se ha observado con razón que los progresos tecnológicos en la esfera productiva hacen que la visión de Marx sea un disparate. Pero ese disparate muestra que Marx no supo separar lo suficiente su visión de las condiciones de producción y no advirtió que esa sociedad que estaba describiendo era Utopía, un lugar cuya sociedad no tendría necesidad de regular la producción general porque la escasez estaría plenamente superada; los seres humanos no tendrían que dedicarse a ninguna actividad productiva genuinamente necesaria. En tales circunstancias una persona podría cazar, pescar, arrear el ganado y criticar según sus ganas. Y en realidad, una persona debe poder hacer todas esas cosas; en Utopía no puede haber escasez, ni siquiera en la forma de autorrealización. Cualquier forma de vida debe ser plena e igualmente accesible para todos.

Parece evidente que esto es por completo imposible. Aun cuando supongamos que se puede superar enteramente la escasez material, siempre quedaría alguna escasez en las formas de la realización humana. Ninguna persona puede realizar plenamente todos los modos posibles de actividad humana. No puede ser un concertista de piano por la mañana, un físico nuclear por la tarde, un neurocirujano al atardecer y un novelista después de cenar, porque el talento y el esfuerzo que exige cada una de esas actividades constituyen el corazón mismo de una vida completa y no dejan lugar a ningún otro. Un hombre puede ser todos los hombres y mujeres, en todo caso como diletante (y ¿para qué sirve ser un neurocirujano diletante?). Pero, actuar como un diletante excluye el compromiso real y pleno que exige la autorrealización.

La escasez en las formas de la realización humana parece un rasgo fijo de la vida humana. Cada uno de nosotros sólo puede realizar por sí una de las muchas vidas posibles que en su conjunto constituyen el florecimiento humano. Aun cuando separáramos cada una de esas vidas de su objetivo instrumental, suponiendo que pudiéramos sentirnos satisfechos realizando actividades que sólo consideramos un juego, comprobaríamos que la escasez se repite en las variedades de los juegos que un individuo puede dominar.

Marx vio, más claramente de lo que con frecuencia se advierte, lo que podía implicar Utopía. Vio que en Utopía nadie tendría que hacer verdaderas elecciones entre las distintas actividades o formas de vida. La "elección" sería un mero artificio para mantener el orden temporal. Elegir la caza por la

mañana no significaría elegir no pescar, puesto que esa otra actividad podría realizarse por la tarde. Entonces Utopía es imposible, porque no podemos escapar a la necesidad de hacer verdaderas elecciones. Supongamos que todos los problemas de producción quedaran resueltos. Si los miedos del Saltamontes fueran fundados, las actividades humanas perderían pues su objetivo. Pero supongamos que podemos alejar por el momento esos temores. Supongamos que la escasez productiva no es una condición necesaria de las actividades humanamente realizables. Quedaría sin embargo la escasez del consumo. Cada uno de nosotros tiene una sola vida que llevar y hay muchas vidas posibles e incompatibles por llevar.

La escasez es el mal humanamente necesario. Ante la escasez, la elección como artificio destinado a seleccionar ciertas actividades y descartar otras es necesaria y la regla de la elección racional es satisfacer al máximo las propias preferencias. Es errado suponer que el hecho de concebir al ser humano como un maximizador de la utilidad esperada es concebirlo de una manera limitada, determinada por las condiciones particulares de nuestra propia sociedad occidental, que no es ni generalmente aplicable ni idealmente apropiada. Es ilusoria una Utopía en la cual la maximización no es expresión de la racionalidad humana. Sólo un dios, al que no podemos concebir, podría morar en semejante reino.

Pero el hombre económico no es por lo tanto el hombre ideal. Porque su no-tuismo es tan inaplicable a la condición humana como la idea utópica de plenitud. Enfrentado a la escasez, pero consciente de que hay otros como él, un ser racional ya no puede dejar de percibir la necesidad de participar, así como no puede dejar de percibir la necesidad de perseguir sus fines y esforzarse. Y el mismo valor intrínseco que encuentra el individuo en la búsqueda y el esfuerzo, lo encuentra también y quizá más plenamente, en la participación. Porque la participación ordena no sólo la escasez productiva, sino también la escasez del consumo.

Apelamos aquí a la concepción de la naturaleza social de los seres humanos sostenida por Rawls quien reconoce su deuda con Wilhelm von Humboldt.[8] Observando que un individuo podía evitar la unilateralidad de su propio desarrollo combinando

armoniosamente sus actividades a lo largo del tiempo, von Humboldt decía: "Lo que en el caso del individuo se logra uniendo el pasado, el futuro y el presente, en la sociedad lo produce la cooperación mutua de sus diferentes miembros; porque, en todas las etapas de su vida, cada individuo puede alcanzar sólo una de esas perfecciones, que representan los rasgos posibles del carácter humano. Por lo tanto, es mediante una unión social, basada en los deseos y las capacidades de sus miembros, que cada individuo puede participar de los ricos recursos colectivos de todos los demás".[9] Rawls ofrece una esclarecedora instancia de la unión social pura: "consideremos un grupo de músicos, cada uno de los cuales se ha entrenado para tocar igualmente bien que los otros cualquier instrumento de la orquesta y que, mediante una especie de acuerdo tácito, ha perfeccionado su habilidad en uno de los instrumentos que es el que ha elegido, de forma tal que al ejecutar juntos se pongan de manifiesto los talentos de todos".[10]

Aquí hay dos tendencias de pensamiento que merecen destacarse. La primera es que cada persona, incapaz de realizar todas las excelencias humanas por sí sola, se beneficia directamente de las excelencias realizadas por los demás que son diferentes, aunque no opuestas a la suya. Cada individuo se enriquece con la complementación de tales realizaciones independientemente de las actividades específicas posibles. La segunda es que al realizar excelencias complementarias, las personas posibilitan actividades que ningún individuo podría realizar por sí solo: las actividades de una orquesta por ejemplo, en oposición a las actividades de todos los músicos separados. En este caso la participación se ejerce más plenamente, y podemos suponer, aun sin calificar la suposición como no-tuista, que cada uno valora las actividades compartidas y la participación. Pero yendo más allá, podemos suponer que al valorar la participación, una persona llega a valorar a los demás participantes, de modo tal que las actividades compartidas crean lazos entre las personas que terminan por hacer que cada uno se interese en los intereses de los otros; aunque, seguramente, no en los intereses de todos los demás, sino de aquellos identificados y experimentados como coparticipantes.

Ahora bien, conviene observar que el valor de la participa-

ción se relaciona con aquello que identificamos como escasez del consumo, la incapacidad que tiene cada individuo de alcanzar todas las formas de realización humana. Pero, en la medida en que algunas de esas formas exigen que se desarrollen actividades compartidas —por más que un individuo domine todos los instrumentos no puede transformarse él solo en una orquesta (dejemos de lado la maravilla tecnológica que permite que mediante registros sucesivos un individuo semejante pueda simular la ejecución de toda una orquesta)—, el valor de la participación revela una insuficiencia del individuo humano que no se reduce a la escasez. No sólo ocurre que un individuo no pueda realizar todas las formas de vida humana por sí solo, también existen algunas formas que no son individualmente realizables.

En 10.1.1 dijimos que la insuficiencia individual era un defecto; la persona consciente de su insuficiencia es consciente de que es menos de lo que quisiera ser. Ahora debemos reconsiderar esa declaración. Dando por sentado que la escasez es un objetivo instrumental dado para nuestras actividades, ¿tenemos razones para lamentar que necesariamente haya que compartir algunas de tales actividades? ¿Tenemos razones para lamentar que ciertas piezas musicales exijan la ejecución de toda una orquesta o que algunos juegos (puesto que no tenemos razón para negar el valor que tienen los juegos en un mundo que no es todo juego) deban realizarse en equipo? En realidad, ¿tenemos razones para considerar que el individuo es defectuoso porque la crianza de los niños parece mejor realizada mediante la cooperación de la madre y el padre? ¿No podríamos suponer mejor que la capacidad social que tenemos de descubrirle valor a la participación es una de las fuentes principales de enriquecimiento de la vida humana, que hace posible (como realmente lo hace) la realización complementaria de nuestras variadas fuerzas y capacidades humanas?

Al valorar la participación, apreciamos la importancia de la vida pública. Nuestra concepción de la sociedad abarca tanto el forum como la plaza del mercado. Pero, así como el valor de la búsqueda y el esfuerzo exige que esas actividades tengan objetivos más allá de ellas mismas, el valor de la participación exige que la vida pública obtenga una base racional de las

variadas realizaciones de individualidad que hace posible. La vida pública, considerada independientemente de su función instrumental podría llegar a ser sin embargo otro juego.

1.3 La moral adquiere un matiz diferente cuando se la considera desde el punto de vista de la relación que tiene con la participación. Para los buscadores y trepadores asociales, la moral puede ser solamente una restricción necesaria pero indeseada. Pero para aquellos que valoran la participación, una moral de acuerdo, aunque sea una fuente de restricción, hace que la actividad compartida sea algo mutuamente aceptado y deseado y por lo tanto estable, que asegura la ausencia de la coerción o del engaño. Como lo mostró nuestro análisis de la cooperación, la restricción acordada es una condición para que los actores acepten racionalmente la división de los beneficios obtenidos mediante la actividad compartida. Y por lo tanto es una condición para que los actores consideren intrínsecamente valorable el hecho de participar en tal actividad.

Una persona valora el esfuerzo, pero sólo, dijimos, en la medida en que considere que los costos de tal esfuerzo son necesarios para alcanzar cierto fin valorado. Si la persona en cambio cree que tales costos no cumplen con la norma de eficiencia implícita en la maximización, ha de considerar mal encaminado su esfuerzo. De manera similar, una persona sólo valora la participación en la medida en que considere que los costos que implica el hecho de participar son necesarios para alcanzar cierto fin valorado. En este último caso no basta con que los costos cumplan la norma de eficiencia implícita en la optimización. Porque si la persona cree que los costos son relativamente mayores que los que debe afrontar cualquier individuo, ha de considerarlos pues excesivos y por lo tanto obtendrá una satisfacción menor de la participación. Por supuesto, puede ocurrir que la persona valore la participación si los costos son excesivos para todos, pero al mismo tiempo cada participante considera intrínsecamente valiosa la actividad compartida. Pues en ese caso la participación cumple la norma de justicia establecida por la concesión relativa minimáxima. Y esto equivale a decir que tal participación debe ser voluntariamente aceptable; los actores deben estar dispuestos a participar sin que se ejerza sobre ellos coerción y sin que se los engañe.

Si las personas desarrollan valores tuistas primariamente en el contexto de las actividades compartidas, esos valores se basan pues en la aceptación de las restricciones morales. Si no hay una división moralmente aceptable de los beneficios, las personas han de considerar que sus actividades compartidas son explotadoras antes que cooperativas y que sus semejantes tienen intereses opuestos a los suyos. Y esta visión es enemiga del desarrollo o del mantenimiento de los lazos tuistas. Podría parecer que, al admitir que las personas llegan a tener interés por los intereses de sus coparticipantes, socavamos la explicación racional de que las restricciones morales tienen una base no-tuista. Pero si las restricciones morales están en la base de los valores tuistas, tales restricciones deben tener un fundamento independiente de esos valores. Las personas llegan a sentir interés por sus semejantes porque reconocen la voluntad mutua de no sacarse ventaja unas a otras y de compartir los beneficios producidos conjuntamente sobre bases justas. Al aceptar las restricciones morales los individuos no expresan su interés por los demás, sino que antes bien crean las condiciones que alimentan tal interés.

Pero al valorar tanto la participación como a los demás participantes, las personas establecen un nuevo valor en la estructura moral dentro de la cual florece la participación. Después de haber conseguido el apoyo racional, la moral obtiene el apoyo afectivo. Las personas demuestran así tener una capacidad afectiva para la moral de la que, según vimos, carece el hombre económico. En 10.2.5 observamos que una capacidad afectiva para la moral no determina las restricciones morales pero supone su reconocimiento previo; el deseo de cumplir con el deber no puede determinar el contenido del deber. Y esto se adapta a nuestro argumento. Las personas reconocen racionalmente las restricciones de la moral como condiciones de una cooperación mutuamente beneficiosa. De modo que llegan a valorar la participación en actividades cooperativas y compartidas que acaten tales restricciones y llegan a interesarse por los demás participantes. Y finalmente llegan a valorar la moral, que al principio sólo les parecía una restricción racional.

Lo que nos interesa aquí no es la psicología moral empírica. No tratamos de describir el proceso mediante el cual los individuos reales llegan a manifestar una capacidad afectiva para la moral. No se logra que los niños se conviertan en seres morales apelando primero a sus intelectos y sólo más tarde a sus sentimientos. Antes bien estamos comprometidos en la tarea del filósofo de la reconstrucción racional. No supusimos que las restricciones morales reales representan el resultado del acuerdo real, pero sostuvimos que, si hemos de justificar las restricciones morales, debemos considerarlas como objetos de un hipotético acuerdo *ex ante*, cuya racionalidad reconocemos ahora *ex post*. De manera similar, no suponemos que los sentimientos morales reales representan el resultado de una valoración previa de la participación y la conciencia de que la participación voluntaria exige aceptar restricciones morales. Antes bien sostenemos que, si debemos considerar que nuestros afectos morales son algo más que sentimientos disfuncionales, de los que nos convendría librarnos, tenemos que poder demostrar que tales afectos morales bien pueden surgir de una valoración y una conciencia semejantes. Debemos mostrar que una persona, al reflexionar sobre sus sentimientos morales, habría de considerarlos una extensión adecuada de su interés por los demás en el contexto de las actividades participativas valoradas, y habría de considerar por lo tanto que es apropiado valorar tales actividades en la medida en que éstas estén moralmente restringidas. Una persona semejante comprobaría que su perspectiva moral es coherente con su reflexión. Si la moral desempeñara en su vida sólo el papel de un mal necesario o de un engaño conveniente, como ocurre en el caso del hombre económico, deberíamos pues esperar que la conciencia de ese papel socavara sus afectos morales.

Una moral racional es contractual. Pero esto no implica que sólo tenga para nosotros un valor puramente instrumental. Al relacionar la moral con la provisión de beneficios que no incluyen en sí mismos ningún interés afectivo por los demás, no disminuimos los sentimientos morales de las personas que sienten tal interés. Precisamente porque podemos darle a la moral una base racional, podemos asegurar también su sostén afectivo.

2.1 Una sociedad justa no es ni una Utopía ni una sociedad de hombres económicos. Una sociedad justa proporciona una

estructura para la comunidad pero no es comunal. La estructura que les brinda a sus miembros promueve la realización de la autonomía de éstos. La división del trabajo que apoya tal sociedad aumenta las oportunidades individuales y alienta la realización de una cantidad más amplia de potencialidades humanas complementarias. Es una sociedad de individuos liberales, libres de establecer sus propios objetivos y de elegir sus propios vínculos afectivos con sus semejantes.

Podemos concebir la relación entre el individuo liberal y la sociedad justa como la que se expresa en la elección arquimédica. Desde el punto de vista arquimédico, la persona racional elige prácticas e instituciones sociales de un modo instrumental, que le permite sin embargo alcanzar la generalidad plena, porque la persona debe identificarse no con cierto individuo particular, o con los miembros de cierto conjunto limitado de individuos, sino con cualquier persona, que podría llegar a ser él en la sociedad elegida. La elección de esa persona refleja, no las preferencias de un individuo o grupo de individuos particular, sino las posibles preferencias de los seres humanos, dadas solamente las relaciones entre las prácticas sociales y los intereses humanos que pueden suponerse racionalmente y las restricciones generales impuestas a las circunstancias establecidas por el horizonte de instituciones y tecnologías factibles.

Una sociedad así elegida es *esencialmente justa*. Sus instituciones y prácticas están destinadas a permitir la maximización del beneficio relativo mínimo, no sólo para un conjunto de personas reales, cuyas preferencias y capacidades se consideran algo dado, ni para personas cuyas preferencias y capacidades están condicionadas socialmente de modo tal que se adapten a lo que ofrecen las instituciones, sino para todas las personas que participan de la interacción para ventaja mutua. Podemos decir que una negociación o un acuerdo particular es instrumental sólo en la medida en que proporcione un beneficio relativo maximínimo a los negociadores reales. Y en 10.2.3 dijimos que un acuerdo es artificialmente justo en la medida en que las preferencias de aquellos a quienes afecta ese acuerdo estén modeladas de manera tal que reciban un beneficio relativo maximínimo. Pero ni la justicia instrumental ni la justicia artificial son suficientes para una sociedad concebida como el

objeto de la elección arquimédica; ambas limitan el número de personas que reciben justicia. La idea de una justicia esencial es que esa cantidad no esté limitada, de modo tal que las instituciones y prácticas se adapten a todas las personas, sean cuales fueren sus preferencias y sus capacidades, con la única condición que todas las personas puedan participar del esfuerzo cooperativo.

Podemos pensar que el mercado perfectamente competitivo, ante el impacto de funciones de utilidad y de producción cambiantes, está continuamente adaptándose en la dirección de un equilibrio que sólo puede lograrse realmente en condiciones estáticas, pero que en la práctica sirve como un ideal regulador que determina la dirección del movimiento del mercado. Del mismo modo podemos pensar que la sociedad esencialmente justa, también ante el impacto de funciones de utilidad y de producción cambiantes, está ajustándose continuamente a fin de hacer posible el beneficio relativo maximínimo. También en este caso, esa posición ideal sólo puede alcanzarse en condiciones estáticas; en la práctica sirve como un ideal regulador que determina la dirección del ajuste social. La neutralidad que ofrece el mercado respecto de los fines de los competidores del mercado se manifiesta en la dirección de los ajustes de tales competidores; la neutralidad se adapta sean cuales fueren los fines de los competidores y hace que la satisfacción de éstos esté en equilibrio óptimo. De manera similar, la neutralidad que ofrece la sociedad esencialmente justa respecto de los objetivos de sus miembros se manifiesta en su adaptación, sean cuales fueren esos objetivos, y hace que la satisfacción de éstos alcance la igualdad relativa óptima.

Una sociedad esencialmente justa es plenamente compatible con la racionalidad crítica y reflexiva de sus miembros. Para brindarles justicia no tiene necesidad de modelar a sus individuos. Ahora veremos cómo la sociedad esencialmente justa obra como un agente socializador; lo que nos interesa aquí es sólo señalar que la neutralidad de la justicia, respecto de los objetivos y preferencias individuales, permite que el defensor de la justicia esencial refute el cargo de que la sociedad suministra un molde al que los individuos deben adaptarse.

Al decir que una sociedad esencialmente justa es neutral

respecto de los objetivos que persiguen sus miembros, negamos que la justicia esté vinculada con cualquier concepción sustantiva de qué es bueno, ni para el individuo ni para la sociedad. Una sociedad justa no tiene otro objetivo que aquel determinado por las preferencias de sus miembros. En su condición de empresa cooperativa destinada a la ventaja mutua, una sociedad justa permite que cada individuo promueva aquello que la sociedad considera bueno. Una sociedad esencialmente justa no presenta un bien social como una función de los bienes individuales. Y esto es lo que más la distingue de una sociedad utilitarista, que también rechaza cualquier concepción sustantiva del bien o cualquier objetivo diferente de aquel basado en las preferencias de sus miembros, pero que supone la existencia de un bien social, que es la suma de los bienes individuales. A una sociedad justa lo único que le importa es permitir que cada persona haga realidad la mayor cantidad posible de su propio bien, en términos aceptados por todos.

Una sociedad esencialmente justa les brinda a sus miembros la oportunidad de gozar del valor intrínseco de la participación. Pero lo hace, no imponiendo estructuras participativas, sino liberando a las personas de los obstáculos a la interacción cooperativa plenamente voluntaria. En realidad, hemos sostenido que las personas racionales deberían aceptar el mercado perfectamente competitivo si se dieran las condiciones para la competencia perfecta o casi perfecta. Pero una sociedad esencialmente justa no les impone el mercado a sus miembros; sin embargo quita lo que podrían constituir barreras al mercado, tanto fortaleciendo la estipulación como castigando el uso de la fuerza y el fraude y también rechazando las prácticas e instituciones sociales compulsivas que encarnen cualquier objetivo sustantivo. Una sociedad esencialmente justa no puede ni prohibir ni exigir que los capitalistas actúen entre adultos condescendientes.

La justicia esencial es una idealización; inventar instituciones y prácticas que se ajustaran perfectamente a los cambios de preferencias, capacidades y circunstancias, que adaptaran a las personas a los cambios y los progresos tecnológicos, excede los poderes humanos. Esto no debe considerarse una defensa de la desesperanza. Los logros obtenidos por las medidas políticas

liberales y por la economía de mercado en el sentido de dirigir a la humanidad hacia la justicia esencial no son despreciables. Pero si una moral basada en el acuerdo apuntala la sociedad esencialmente justa y si una teoría de la economía basada en el acuerdo (que evidentemente es la teoría del mercado) trata la provisión de bienes privados en una sociedad semejante, la selección y provisión de bienes públicos en tal sociedad exige una teoría de una política basada en el acuerdo y no hay duda de que esa teoría aun no ha sido desarrollada. Los teóricos de la elección pública[11] han ofrecido significativas contribuciones en este sentido, pero no podemos considerar aquí los problemas del procedimiento político ni del diseño institucional.

2.2. Desde el punto de vista arquimédico, un individuo racional debería elegir la justicia esencial. En 10.2.3 sostuvimos que la justicia artificial, al adaptar las personas a las instituciones, no logra acomodar adecuadamente la capacidad racional de las personas para reflexionar de manera autocrítica sobre sus preferencias. La elección arquimédica acomoda esa capacidad. Aun cuando una persona, consciente de sus utilidades y capacidades, no se encuentre en el punto arquimédico, puede reconocer la importancia que tiene esa posición para aquellas elecciones que determinarían el contexto social general de su propia actividad, porque esa posición le asegura que no sólo sus preferencias actuales, sino también aquellas preferencias posibles que puede llegar a tener en el futuro, serán justamente tenidas en cuenta.

El individuo racional debe afrontar una incertidumbre fundamental, tanto al comprenderse a sí mismo como al tratar sus interacciones con los demás. Esa incertidumbre fundamental surge de la dimensión reflexiva de la racionalidad. La función de utilidad que define parcialmente a un individuo particular puede ser revisada, en parte mediante la propia actividad reflexiva del individuo. Ahora bien, hay cuestiones profundas y difíciles relativas a la identidad personal que están vinculadas con ese carácter revisable. Derek Parfit examinó muchas de esas cuestiones al considerar la relación que existe entre los diferentes "sí mismos" temporalmente distantes.[12] Breve y dogmáticamente presentamos la siguiente posición. La identidad personal tiene una dimensión esencialmente prácti-

ca. El sí mismo en el momento *t1* es idéntico al último sí mismo del momento *t2* en la medida en que se identifica con ese último sí mismo, y esa identificación se mide por el valor dado a las preferencias esperadas del sí mismo del *t2*, en la preferencia del sí mismo del *t1*. La utilidad es, como hemos insistido permanentemente, la medida de la preferencia presente, las preferencias del sí mismo en un momento particular; la racionalidad práctica es la maximización de la utilidad y por lo tanto la maximización de la satisfacción de las preferencias presentes. Las preferencias presentes pueden o no acomodarse a las preferencias futuras anticipadas; en principio ésta no es una cuestión de racionalidad. La persona racional puede ser una persona prudente, aunque no necesariamente. Sin embargo, la naturaleza autocrítica de la racionalidad agrega una dimensión adicional de gran importancia para determinar la relación existente entre los diferentes sí mismos distantes en el tiempo. Porque la persona plenamente racional debe considerar que sus sí mismos futuros se vinculan con su sí mismo presente mediante un proceso de reflexión crítica; las preferencias futuras surgen de la preferencia presente mediante un ejercicio racional, por supuesto basado en la experiencia. La conexión que existe entre el sí mismo del *t1* y el sí mismo del *t2*, no es por consiguiente ni arbitraria ni enteramente accidental. Y según creemos, esto dispone a la persona racional a identificarse con sus sí mismos posteriores, dándoles a aquellas preferencias futuras menor valor que a sus preferencias presentes, pues espera que aquéllas deriven de éstas mediante un proceso racional.

Pero esa identificación con los sí mismos futuros implica incertidumbre. La persona racional ignora parcialmente las preferencias de ese sí mismo temporalmente extendido a quien le importa satisfacer. Ahora bien, en muchas de sus elecciones esa ignorancia no tiene un efecto práctico. Pero al estar de acuerdo con las instituciones y prácticas de la sociedad —su medio social general—, el individuo debe acomodar esa ignorancia. Y esto lo induce a acercarse al punto arquimédico al elegir entre las diferentes estructuras sociales. Porque el interés de esa persona por sus posibles sí mismos futuros se transforma, en efecto en un interés por ser justo con aquel que resulte ser y el punto arquimédico le asegura esa imparcialidad para todos

sus posibles sí mismos. Aunque ningún individuo necesita ni puede suponer que esos sí mismos futuros sean todas las personas posibles, la dificultad de decidir el propio futuro hace que el punto arquimédico constituya un artificio útil para expresar la identificación del individuo con cualquier preferencia que pueda llegar a tener después.

La capacidad de reflexionar de manera racional y autocrítica sobre los propios fundamentos de la acción puede identificarse con la idea tradicional de autonomía. La persona que se considera su propio juez no está atada a sus preferencias presentes, salvo en el caso de que éstas constituyan su punto de partida para la autoevaluación. En ausencia de cualquier tipo de preferencias y capacidades, el individuo carecería de una base para considerar qué clase de persona sería. Aunque aceptamos la idea kantiana de la autonomía, continuamos rechazando la afirmación kantiana de que la razón proporciona fundamentos sustantivos para la elección o la acción.[13] Pero reconocemos que la razón ofrece una base formal para la elección arquimédica, congruente con la autonomía que confiere la racionalidad. De modo tal que la sociedad esencialmente justa es el hábitat adecuado del actor plenamente autónomo. La persona a la cual le importa el ejercicio pleno de su libertad racional no puede estar de acuerdo con instituciones y prácticas que sean sólo instrumentalmente justas, por muy adaptadas que estén tales instituciones y prácticas a sus intereses y poderes presentes, porque el individuo no tiene ninguna garantía que que tal estructura social continúe satisfaciéndolo con justicia en el futuro.

La segunda incertidumbre que deben afrontar los individuos racionales tiene que ver con la disposición de los otros a aceptar voluntariamente las instituciones y prácticas sociales. Y dado que, como sostuvimos, tal disposición se basa en la justicia de la sociedad, un individuo sólo puede estar seguro del acatamiento voluntario en una sociedad esencialmente justa. Porque de lo contrario, los cambios que se produzcan en las preferencias, los poderes o las circunstancias de las demás personas pueden hacer que los acuerdos sociales se vuelvan injustos y pueden socavar su continuado carácter aceptable. Por lo tanto, el interés por la estabilidad también produce un

movimiento hacia el punto arquimédico en el momento de elegir entre las estructuras sociales. También en este caso podemos decir que ningún individuo necesita suponer, ni puede suponer, que ha de encontrarse en interacción con todas las personas posibles o en todas las circunstancias posibles. Sin embargo, la dificultad para encontrar los límites plausibles de la cantidad tanto de las personas como de las circunstancias, especialmente en una sociedad populosa y tecnológicamente avanzada, hace que el punto arquimédico se convierta en un útil artificio que permite asegurar la estabilidad de las instituciones y prácticas que hacen posible y beneficiosa la cooperación.

Aunque las personas racionales que valoran las actividades participativas con sus semejantes adquieran un interés afectivo por la justicia esencial, el objeto de su interés es inicialmente un valor instrumental. Los componentes de una moral racionalmente convenida se juzgan por el modo en que contribuyen a la satisfacción de las preferencias dirigidas a objetos por completo diferentes. Pero el carácter instrumental de la justicia esencial es paralelo al carácter instrumental de la racionalidad. Y ninguna de las dos llega a ser meramente el mejor medio para alcanzar algún fin particular o algunos fines particulares. Antes bien, cada una tiene valor dados ciertos rasgos generales característicos de la actividad humana. La razón práctica sólo le es útil a un individuo que tiene la capacidad y la necesidad de elegir entre varias acciones posibles, pero dadas esa capacidad y esa necesidad, la razón tiene necesariamente valor para tal individuo. Del mismo modo, la justicia sólo le es útil a un individuo que tiene la capacidad y la necesidad de elegir entre varias interacciones posibles en las cuales el valor que tenga el resultado para cada persona depende de las acciones de todos, pero dadas esa capacidad y esa necesidad, la justicia tiene necesariamente valor para ese individuo. En última instancia, los seres humanos valoran la justicia porque ésta les proporciona la base para las interacciones, libres de toda restricción externa, interacciones que les permiten satisfacer mejor las preferencias sean cuales fueren éstas y sean cuales fueren las capacidades de los individuos que participen de tales interacciones.

Glaucón se representa la justicia como un mal necesario

porque supone que la justicia sólo nos proporciona un medio (a falta de otro mejor) para alcanzar nuestros fines. Pero Glaucón se equivoca. La cooperación hecha posible gracias a la justicia, considerada tanto atendiendo a lo que permite lograr como atendiendo a la actividad participativa que implica, no es una segunda forma (a falta de otra mejor) de lograr aquello que, si no fuera por ciertos obstáculos particulares, podría alcanzarse de algún otro modo. Al cooperar, hacemos el uso más efectivo de nuestros poderes para alcanzar los fines que, de otro modo, no podríamos alcanzar con nuestra capacidad individual. Y así encontramos el valor en modos que ningún individuo solitario podría experimentar. Por supuesto es cierto que un ser todopoderoso y autosuficiente no emplearía la cooperación y por lo tanto tampoco su virtud, la justicia. Pero ninguno de nosotros es ni todopoderoso ni autosuficiente. De modo que no podemos suponer razonablemente que la cooperación que aumenta nuestros limitados poderes y supera nuestra insuficiencia individual, es un mal necesario. En realidad, concebir la cooperación, y por lo tanto la justicia, como males necesarios equivale a considerar que los seres humanos son en sí mismo malos.

Aunque Glaucón se equivoque al representarse la justicia, el paralelo que nos señala nos da una importante lección. La justicia es una virtud que se ajusta a los poderes y deseos humanos. El relato del pastor lidio muestra las consecuencias de conferirle el poder de obrar "como un dios" a alguien que conserva los deseos humanos. Las ataduras que impone la justicia pierden su influencia sobre aquellos que pueden gozar de los beneficios sin pagar sus costos. Hacemos pues bien en refrenar a ese dios mortal, el Leviatán, no sea que se imagine que la justicia es un mal innecesario.

2.3 Una moral basada en el acuerdo suma al entendimiento del hombre económico, los afectos del individuo liberal. Las características que definen la individualidad liberal han de encontrarse, explícita o implícitamente, en nuestro análisis anterior. Aquí sólo tenemos que recapitularlas y volver a poner el acento en ellas. Ante todo, el individuo liberal es un ser activo, que obtiene satisfacción en la búsqueda y el esfuerzo que constituye lo que los humanos concebimos como actividad. Por supuesto, para reiterar lo que señalamos en 1.1, si una persona

le da valor intrínseco a una actividad instrumentalmente valiosa, también debe darle valor a la finalidad de esa actividad; la buena vida debe combinar los logros con el esfuerzo. Pero, en la medida en que este último es esencial, el hábitat del individuo liberal debe ser un ambiente de escasez, sea ésta material, mental o emocional.

En segundo lugar, el individuo liberal tiene su propia concepción independiente del bien. Y esa concepción no necesariamente debe apoyarse en una motivación asocial, porque el individuo liberal no es un individuo asocial. Pero característicamente, los bienes de los diferentes individuos reflejan tanto armonías como conflictos de intereses. El bien de cada persona expresa sus preferencias, pero, como ya lo hemos dicho, sus preferencias consideradas. Por consiguiente, el individuo liberal debe tener la capacidad de reflexionar sobre sus preferencias y de alterarlas a la luz de tal reflexión; esa capacidad lo hace autónomo.

En tercer lugar, pues, el individuo liberal es plenamente racional, posee una racionalidad que incluye tanto la autonomía como la capacidad de elegir entre las acciones posible sobre la base de su propia concepción del bien, determinada por sus preferencias reflexivas. Al definir la utilidad como la medida de la preferencia considerada, el individuo es racional si se convierte en un maximizador de la utilidad esperada. Ponemos aquí el acento en la dimensión autocrítica de la racionalidad práctica, puesto que ésta aparentemente carece de la cruda concepción de la racionalidad meramente económica.

Como ser autónomo, el individuo liberal es consciente del proceso reflexivo mediante el cual sus sí mismos futuros han de surgir de su sí mismo presente, de modo tal que sus preferencias se modifican no al azar o de manera descontrolada, sino a la luz de su propia experiencia y su propio entendimiento. Esto le brinda el sentido de un ser extendido en el tiempo, cuya vida es un único desarrollo que puede integrarse y unificarse. Por lo tanto, aunque el individuo liberal no necesariamente debe ser prudente y tener una visión atemporal de sus preferencias y de la satisfacción de tales preferencias, observa el presente no como algo aislado, sino en relación tanto con su pasado como con sus posibles futuros. Y por consiguiente, es capaz de restringir su

conducta maximizadora porque considera que tal restricción es racional. Tiene pues la capacidad práctica para aquello que hemos identificado como conducta moral.

Pero, más allá de esa capacidad práctica, el individuo liberal tiene una capacidad afectiva o emocional. Porque advierte que debe elegir entre varios estilos de vida posibles y que la amplitud y la riqueza de sus elecciones dependen de la existencia de otras personas que eligen otras formas. Por lo tanto, el individuo liberal estima que su vida se desarrolla en un contexto social, que sólo es posible gracias a la interacción con los demás, interacción que, por supuesto, también hace posible las vidas de los otros. Al valorar esa amplitud de oportunidades y la riqueza de la satisfacción, el individuo liberal llega a valorar la participación tanto como la búsqueda y el esfuerzo y, aunque no sea un tuista natural, llega a valorar a aquellos semejantes que participan con él en la interacción. De modo que el hábitat del individuo liberal se caracteriza tanto por la escasez como por la sociedad, y la satisfacción intrínseca que obtiene de la sociedad no está disociada de la función instrumental que cumple ésta para superar la escasez. Al valorar la participación y a sus compañeros participantes, el individuo liberal valora también las restricciones que hacen posible tal participación. El intelecto y la emoción se suman para hacerlo un ser moral.

Pero hay otra característica del individuo liberal que debemos destacar aquí. Nos referimos a la afectividad libre. El individuo liberal no carece de ataduras emocionales con las demás personas, pero los lazos que lo ciñen son voluntarios, o para decirlo más apropiadamnte, representan la voluntad conjunta de las personas a las que ciñen. Así como cada individuo tiene su propia concepción del bien y hace su propia elección entre los estilos de vida posibles, del mismo modo cada individuo elige a otros como objetos de afecto. No está atado por roles sociales fijos, ni en sus actividades ni en sus sentimientos. Aunque para el individuo liberal las relaciones afectivas sociales son esenciales, no hay para él relaciones sociales esenciales.

Ahora bien, no hay que confundir la afectividad libre con el concepto tradicional de "amor libre". Particularmente, no hay razón para que las personas cuyos afectos son libres y están sujetos a su propio control autónomo, no entablen relaciones

duraderas y obligatorias con los otros. Lo esencial de la afectividad libre es que sus vínculos sean los que cada persona quiere crear. En la esfera de las emociones, la imposición, la falta de compromiso, son incompatibles con la autonomía individual.

La afectividad libre está vinculada con la justicia esencial. Sólo en una sociedad cuyos miembros pueden considerarla una empresa cooperativa genuina destinada a la ventaja mutua, podemos esperar que existan relaciones puramente voluntarias entre las personas con la fuerza suficiente para constituir el cemento social. Hay dos razones relacionadas entre sí para que esto sea así. Por un lado, una sociedad a la que se le reconoce la tarea de promover la ventaja mutua de un modo justo necesita menos ataduras personales para conservarse; las personas participan cooperativamente de la interacción por los beneficios impersonales que ésta acarrea. El hecho de que las personas que no se interesan por los intereses de los demás reconozcan a los otros como socios de una empresa mutuamente beneficiosa, les permite participar en la interacción de manera constructiva. Por otro lado, en una sociedad reconocida como una empresa cooperativa justa podemos esperar que florezca una asociación cívica genuina. Los vínculos interpersonales que no podrían existir entre personas que se vieran unas a otras de una forma tradicionalmente competitiva (una forma que distinguimos de la competencia del mercado) o como enemigos, se generan naturalmente en un marco de beneficio mutuo. De modo que una sociedad esencialmente justa alimenta los lazos interpersonales y al mismo tiempo depende menos de ellos para conservarse. Y así, una sociedad esencialmente justa puede brindar la posibilidad a sus miembros de elegir libremente sus vínculos interpersonales.

En ausencia de una capacidad afectiva para la moral entre sus miembros, la afectividad libre demostraría rápidamente que es desestabilizante para cualquier sociedad. A fin de que los individuos respeten voluntariamente las restricciones impuestas por la justicia esencial es necesario que exista tal capacidad afectiva para la moral. Si los miembros de una sociedad no sienten el compromiso emocional de mantener la estructura de beneficio mutuo, esa sociedad sólo puede garantizar su propia estabilidad imponiéndoles a sus miembros, mediante procesos

de socialización, la lealtad a un objetivo más sustantivo, el cual determinaría roles que los individuos no podrían aceptar o rechazar libremente. El amor libre puede haber surgido como una reacción contra la responsabilidad emocional; una reacción comprensible en tanto esa responsabilidad les fue impuesta a las personas independientemente de sus preferencias y sus voluntades. Pero la afectividad libre supone precisamente que la persona es responsable de sus emociones.

La idea de la afectividad libre seguramente ha de parecerles amenazadora a los defensores de la comunidad sustantiva, sean éstos radicales o reaccionarios. Sin duda la preocupación de éstos puede parecer más justificada en la esfera de la sociedad puramente privada o doméstica. Aunque la afectividad libre no excluye de ningún modo el supuesto de que la buena crianza de los niños exige el compromiso emocional, puede suponerse también y con cierto fundamento, que son muy pocos los adultos que pueden prepararse para asumir ese compromiso de un modo elegido libremente. La negociación justa entre las generaciones que analizamos en 9.6 era, como se recordará, una negociación esencialmente pública, relativa a proporciones sociales de inversión. No involucraba demandas emocionales directas. La crianza de los niños, en una sociedad cuya tecnología brinda otras formas de atender a los ancianos, es una cuestión por completo diferente.

La afectividad libre no es un ideal carente de problemas. No podemos suponer que una sociedad de individuos liberales genere intereses intergeneracionales que aseguren su viabilidad a lo largo de los años. Pero igualmente, tampoco podemos decir con seguridad que no los genere. Ese es un juicio que por ahora debemos aplazar.

2.4 Al defender la prioridad normativa del individuo ante la comunidad, no decimos nada respecto de la base causal de la individualidad. La autoconciencia que necesariamente debe tener un ser para tener una autoconcepción genuina, sólo puede resultar de la socialización. Y la autoconciencia está en la raíz de nuestro sentido profundo de individualidad; en la perspectiva del ser autoconsciente, los fundamentos para actuar deben basarse en el propio yo. Al producir un ser autoconsciente, la sociedad humana se cuestiona a sí misma. Desde el punto de

vista del ser autoconsciente, las prácticas e instituciones sociales constituyen la encarnación de normas surgidas de la necesidad que tiene el individuo de justificar tales instituciones y prácticas para poder luego aceptarlas voluntariamente.

Pero si admitimos que la individualidad puede tener una causa social, de modo tal que las personas son productos sociales, ¿no deberíamos pues rechazar el argumento de que las personas son autónomas? Y si sacrificamos la autonomía, ¿no estamos acaso socavando la concepción del individuo liberal, que constituye la médula misma de nuestra respuesta a aquellos que repudian la moral basada en el acuerdo por considerarla una seudomoral del hombre económico? La idea de que la socialización es una amenaza para la autonomía no es nueva, y si supusiéramos que un ser autónomo debe constituirse *ex nihilo* no podríamos refutarla. Pero la autonomía no tiene semejante implicación. Sería absurdo identificar al individuo con el proceso formal de reflexión y de elección en el cual se manifiesta la autonomía. Ese proceso requiere un material —las preferencias y las capacidades— que le sirva de punto de partida, y reconocer que esas contribuciones iniciales no se determinan, por lo menos al principio, de manera autónoma, no constituye una amenaza a la autonomía. Lo que hace a un ser autónomo es su capacidad de alterar las preferencias dadas mediante un procedimiento racional, autocrítico y reflexivo y no una capacidad para producir preferencias sin ninguna base previa.

Podemos suponer que un individuo comienza sus interrelaciones con preferencias y capacidades que, por lo menos en parte, están determinadas socialmente. Y suponemos que la capacidad por la cual un individuo reflexiona de manera autocrítica sobre sus preferencias y capacidades está también en parte determinada socialmente. En efecto, suponemos que los seres humanos están socializados en la autonomía. Lo importante es que ni sus preferencias ni (dentro de ciertos límites) sus capacidades queden fijadas por esa socialización, que no es un proceso mediante el cual se ata firmemente a las personas, sino antes bien un proceso, al menos en parte, que permite el desarrollo de seres atados débilmente que tienen la capacidad de cambiar el modo en el cual están constituidos.

En 2.1 caracterizamos la sociedad esencialmente justa

como una sociedad neutral respecto de los objetivos de sus miembros, que se adapta a ellos con el fin de permitirles obtener un resultado justo y óptimo, sean cuales fueren sus preferencias particulares. Ahora podemos agregar algo a esa caracterización y decir que una sociedad esencialmente justa es también neutral respecto de la socialización de sus miembros, en tanto les asegura la autonomía necesaria para alterar sus preferencias mediante la reflexión racional. El punto arquimédico exige tanto la neutralidad respecto de la producción de las preferencias como la neutralidad respecto de su satisfacción.

No obstante, la neutralidad respecto de la producción de las preferencias debe ser limitada. Suponemos que las personas sólo están atadas débilmente, de modo tal que pueden cambiar sus deseos y sus objetivos. No podemos suponer que el proceso de socialización sea neutral respecto de lo que podemos llamar, siguiendo con la metáfora, el extremo inicial del hilo que ata a los miembros de una sociedad. La sociedad esencialmente justa puede ser neutral respecto de la alteración de las preferencias, pero no puede ser neutral respecto de su producción inicial. El proceso de socialización crea personas determinadas, no la mera posibilidad de que tales personas estén constituidas mediante un no imaginado y quizás inimaginable acto de autocreación. En la actividad socializadora debe haber pues un momento no neutral, hasta en el caso de la sociedad esencialmente justa. Y ese momento no neutral, aun cuando no implique una amenaza a la autonomía, parece ser un factor arbitrario que se basa en la necesidad de justificación del individuo autónomo racional.

Pero, ¿acaso la neutralidad asegurada de la sociedad esencialmente justa no suministra tal justificación? Damos por sentado que cada persona está constituida en parte por una serie inicial de preferencias que no puede estar sujeta a su propio control y que, desde el punto de vista del individuo, es el resultado arbitrario del proceso social mediante el cual, él, como un individuo con determinadas preferencias y capacidades, nació. Sin embargo, si ese individuo es un miembro de una sociedad esencialmente justa, tiene la garantía de que esa arbitrariedad no puede afectar su satisfacción equitativa; sean cuales fueren las preferencias que pueda tener la persona, o que pueda llegar a tener como resultado de reflexionar sobre las

preferencias dadas inicialmente, en una sociedad esencialmente justa ese individuo puede alcanzar un resultado justo y óptimo.

Aunque parezca evidente, digamos que una sociedad esencialmente justa no puede socializar arbitrariamente a sus miembros. En 2.1 señalamos que una sociedad esencialmente justa se relaciona con una elección que refleja las posibles preferencias de los seres humanos, atendiendo solamente a las relaciones generales entre las prácticas sociales y los intereses humanos que pueden suponerse razonablemente, y las restricciones generales a las circunstancias establecidas por el horizonte de factibilidad. Una sociedad esencialmente justa no ha de alentar anhelos por lo imposible; no ha de alentar objetivos radicalmente opuestos que amenacen la perspectiva de la ventaja mutua. Nadie tiene razón para quejarse de esto porque nadie tiene razón para considerarse perjudicado por tales restricciones a la posible socialización.

En realidad, debemos suponer que en una sociedad esencialmente justa, la socialización fortalece las instituciones y prácticas que sus miembros apoyan racionalmente. Mientras deja que cada persona elija libremente los objetos particulares de sus afectos, la sociedad debe alentar la sociabilidad general de esos miembros, puesto que tal sociabilidad facilita la interacción. La garantía de que la interacción está sujeta a las restricciones de la justicia satisface la preocupación enfatizada en el pensamiento feminista de que la socialización no constituya una base para la explotación. En una sociedad esencialmente justa, ningún individuo debe tener que sacrificar sus intereses a los intereses de los demás a fin de obtener la satisfacción de la interacción social.

Además, una sociedad esencialmente justa debe fortalecerse mediante el desarrollo de los afectos e intereses de los jóvenes, de modo tal que los intereses madurados de la juventud proporcionen una motivación que refuerce los requisitos racionales de la cooperación. La actividad cooperativa debería experimentarse como satisfactoria en sí misma. Y también debería hacerlo la actividad de mercado, pero en la medida en que ésta no exige las restricciones a la conducta maximizadora de la utilidad que son indispensables para una cooperación justa, es de esperar que las

satisfacciones no sociales constituyan una motivación más extensa y más efectiva dentro del mercado que fuera de él. La socialización debería pues alentar a las personas a desear cooperar en aquellas situaciones en las cuales la cooperación sólo puede ser mutuamente ventajosa para ellas. El deseo de cooperar en tales circunstancias recibiría pues el apoyo reflexivo de la razón; la justificación de la sociedad esencialmente justa se extiende a la justificación de la sociabilidad que la sostiene y fortalece y, por lo tanto, a la justificación de la socialización que instila y alienta esa sociabilidad.

Precisamente porque una sociedad esencialmente justa hace posible el mayor grado de satisfacción al alentar a las personas a realizar las muchas y variadas potencialidades de los seres humanos, debemos esperar que tanto en las instituciones como en los procedimientos socializantes de la sociedad se instaure la división del trabajo. Semejante división del trabajo no ha de expresarse, por supuesto, en la forma de requisitos que terminen por someter la autonomía individual. No ha de ser coercitiva. Antes bien, deberíamos esperar que una sociedad esencialmente justa promueva una concepción ampliamente pluralista de la realización humana que, dados los accidentes que diferencian circunstancias y capacidades, lleve a los individuos a elegir roles diferentes pero complementarios. Ni la objeción presentada por Marx, de que la división del trabajo impide el desarrollo multilateral de los seres humanos individuales, ni la objeción presentada por Rousseau, de que la división del trabajo hace que los hombres dependan de las voluntades ajenas de los demás y por lo tanto les quita libertad, pueden sostenerse en las circunstancias de la justicia esencial.[14] Porque, como ya vimos antes, ningún ser humano es capaz de realizar todas las formas posibles de la actividad humana; la forma más completa posible de realización para cada uno de nosotros es la realización de un modo que se complemente con los demás modos que realizan los otros. Y en la sociedad esencialmente justa, nadie resulta explotado o está en desventaja por depender de los demás; la insuficiencia de cada individuo, tanto en lo relativo a la producción de los bienes materiales de la vida humana, como en lo relativo a la realización de los variados modos de vida humana, se supera de la manera más

efectiva compartiendo las actividades justas y óptimas que tal sociedad hace posibles.

Pero hay aspectos más problemáticos referentes a la socialización que parecen amenazar la posibilidad misma de una sociedad esencialmente justa. Todas las sociedades humanas conocidas apoyan la diferenciación de los sexos y la diferenciación de status, y esas dos diferenciaciones están vinculadas entre sí; los varones tienen un status consistentemente superior al de las mujeres.[15] Ahora bien, el contenido de la diferenciación por sexos y de la diferenciación por status es en alto grado (aunque no totalmente) variable; lo que se mantiene constante es la forma de diferenciación. Pero, dada esa constancia, tendría que ser muy osada la persona que supusiera que es posible eliminar tanto la diferenciación por género como su vínculo con el status. Aun cuando supusiéramos que la diferenciación por sexos, la diferenciación e identificación conceptuales de los seres humanos correspondiente a las diferencias fisiológicas y sexuales, es en sí misma un producto enteramente social, ésta podría ser el resultado de una diferencia fuertemente condicionada en las respuestas de conducta de los seres humanos adultos a las diferencias sexuales de la infancia y la niñez. El hecho de tratar a los niños como varones y a las niñas como mujeres y de socializarlos de manera tal que se vean a sí mismos como esencialmente varones o mujeres, puede ser algo fijo.

La diferenciación por sexos amenaza en sí misma la justicia esencial. El efecto que provoca es relacionar las expectativas de las mujeres y de los varones con ideales separados o con determinados modelos. Las instituciones y prácticas sociales, alientan (aunque en realidad no las exijan) diferencias de roles entre las mujeres y los hombres. Y esas diferencias están vinculadas con las diferencias afectivas; la socialización promueve diferentes formas de sociabilidad para hombres y mujeres. Aun cuando ninguna de esas diferencias estuviera relacionada con el status, el hecho de imponer la diferenciación por sexos al determinar la satisfacción adecuada para hombres y mujeres, parecería incoherente con los requisitos de la elección arquimédica que determinan la justicia esencial.

El status diferencial es una cuestión aun más clara. No puede suponerse que una sociedad que lo atribuye sobre la base

de normas impuestas cumpla con la neutralidad de la satisfacción que es un requisito de la justicia esencial. Pero si la asignación del status diferencial es uno de los puntos fijos de la sociedad humana, la justicia esencial es imposible. Puede parecer pues que el momento no neutral de la socialización socava la idea de una sociedad de individuos autónomos, libres afectivamente y liberales.

3 Podemos interpretar que la primera indagación realizada sobre la justicia, la *República* de Platón, constituye un intento de defender la *polis* griega, representando su naturaleza ideal y mostrando que esa naturaleza es congruente con la razón. Podríamos decir que nuestra indagación también se basa en un interés semejante y que constituye un intento de defender la sociedad occidental de mercado, vinculando su naturaleza ideal con la razón. Pero así como la *polis* verdadera le quedaba irremediablemente chica a la ciudad descrita en palabras, la sociedad de mercado le queda chica a la sociedad esencialmente justa.[16] Sin duda, Platón abrigaba la esperanza de transformar la *polis* en su ideal, pero su articulación llegó demasiado tarde. ¿También aquí hay un paralelo?

Al finalizar *El orden de las cosas*, Michel Foucault escribe: "La arqueología de nuestro pensamiento muestra fácilmente que el hombre es una invención de reciente data. Y quizás una invención que está próxima a su fin".[17] Ya sea una invención, ya sea un descubrimiento, el individuo liberal es en realidad un ser de reciente data; ¿no será quizá la mera criatura de un momento histórico, no serán acaso su racionalidad maximizadora y su afectividad libre variaciones de un tema más amplio? Hemos comprobado que su nicho ecológico está en la sociedad esencialmente justa que en su sentido más acabado es "una empresa cooperativa destinada a la ventaja mutua". Hemos comprobado que la moral y la sociabilidad del individuo liberal surgen de (y se sostienen gracias a) considerarse a sí mismo como un participante de la interacción con los demás en modos que le proporcionan a todos beneficios mutuos justos. En otros ambientes carentes de la justicia esencial, las personas estarían restringidas, no de manera autónoma por la razón, sino de manera heterónoma por creencias y afectos condicionados socialmente. De modo que podemos preguntarnos respecto de las perspecti-

vas que tiene el individuo liberal: ¿su nicho ecológico puede llegar a convertirse en realidad y a conservarse?

Históricamente, el individuo apareció, tanto desde el punto de vista práctico como desde el punto de vista conceptual, en un ambiente en el cual hubo una fuerte correlación positiva entre la prosecución del interés individualizado y un aumento continuado de la producción de bienes y servicios para todos. No obstante, por justa o injusta que pueda ser la distribución de los beneficios, la sociedad occidental ha mantenido la apariencia de proporcionar a sus miembros una ventaja mutua. Y podemos atribuirles papeles muy significativos en la creación de ese ambiente a la perfección de los instrumentos sociales destinados al intercambio y la inversión, y también al desarrollo de una tecnología orientada a aumentar la producción y la diversificación. Pero carecemos de la pericia necesaria para formarnos una expectativa razonable respecto de las perspectivas de continuidad y mucho menos de mejoramiento de esas circunstancias favorables a la individualidad liberal.

Igualmente carecemos de la habilidad para formarnos cualquier expectativa razonable sobre el futuro conceptual del individuo. Si nuestra autocomprensión actual es el producto de la desaparición de la creencia de que el hombre es una criatura que recibe sus objetivos y sus valores de un orden divinamente establecido, bien podría ser que esa autoconcepción no fuera más que una incómoda etapa de transición hacia la aparición de la creencia de que el hombre es un artefacto, construido en virtud de un conocimiento creciente y un proceso alterable de la ingeniería genética. Y ese hombre producto de la ingeniería quizá no sea más compatible con las ideas de individualidad y autonomía que están en la base de nuestro argumento, que el hombre-criatura.

Quizá lo que hemos construido no sea una teoría que vincula la moral con la elección racional, sino una representación de las restricciones morales y de la elección maximizadora como se dan en una efímera sociedad de mercado. Sin embargo, teníamos la esperanza de llegar a una mejor conclusión. Nietzsche comienza el segundo ensayo de *La genealogía de la moral* con esta pregunta: "Criar un animal *con el derecho a hacer promesas*, ¿no es ésa la paradójica tarea que se ha impuesto la

naturaleza en el caso del hombre? ¿No es acaso ése el principal problema relativo al hombre?"[18] La naturaleza no se impone ninguna tarea, pero en la metáfora de Nietzsche podemos encontrar la médula de la autocomprensión que, una vez alcanzada no se puede sacrificar con liviandad. Un animal con el derecho a hacer promesas debe ser capaz de comprometerse, de darse una razón para decidir y obrar, de modo que superen su interés habitual por satisfacer sus preferencias. Un animal semejante es capaz de participar en la interacción de su mundo de un modo nuevo y distintivo, que creemos haber captado en la concepción de la maximización restringida. El hombre económico podría ser sólo "una máquina globalmente maximizadora".[19] Pero el individuo liberal seguramente es "consciente de su superioridad sobre todos aquellos que carecen del derecho a hacer promesas y se consideran sus propios garantes... y es consciente de hasta qué punto su dominio sobre sí mismo también le brinda necesariamente el dominio sobre las circunstancias, sobre la naturaleza y sobre todas las criaturas con menos voluntad y ninguna confiabilidad".[20] Y ese dominio sobre sí mismo se expresa cuando, al reconocer la necesidad de una comunidad, el ser humano individual, varón o mujer entre varones y mujeres, se adhiere a una moral basada en el acuerdo.

Notas

1 B. Suits, *The Grasshopper: Games, Life and Utopia*, Toronto, 1978, pág. 170.
2 *Ibid.*, pág. 167.
3 *Ibid.*, pág. 41.
4 *Ibid.*, pág. 176.
5 *Ibid.*, pág. 177.
6 Véase Aristóteles, *Etica a Nicómaco*, X, caps. 6-7; *Metafísica*, XII, caps. 7, 9.
7 Karl Marx y Frederick Engels, *The German Ideology*, p.1, comp. C. J. Arthur, Nueva York, 1970, pág. 53.
8 J. Rawls, *A Theory of Justice*, Cambridge, Mass., 1971, págs. 522-527.
9 W. von Humboldt, *The Limits of State Action*, J. W. Burrow, (comp.), Cambridge, 1969, págs. 16-17.
10 Rawls, *op. cit.*, pág. 524 n.

11 En particular J. M. Buchanan, a cuyo libro *The Limits of Liberty* nos referimos en el capítulo 7.

12 Véase D. Parfit, "Later Selves and Moral Principles"; también "Personal Identity", *Philosophical Review* 80, 1971, págs. 3-27.

13 Podría negarse que Kant haya hecho tal afirmación. Pero nosotros consideramos que está implícita en la idea del imperativo categórico.

14 Ya citamos el punto de vista de Marx en 1.2. En el caso de Rousseau, véase la nota 31 del capítulo 10. La idea que subyace en Rousseau está claramente formulada en *Projet de constitution pour la corse*, en *Political Writings*, vol.2, pág. 308: "Cualquiera que dependa de otro y que carezca de los recursos en sí mismo, no puede ser libre".

15 Véase, por ejemplo, Ernestine Friedl, *Women and Men: An Anthropologist's View*, Nueva York, 1975; "Comenzamos con la prueba de que en todas las sociedades conocidas existe algún grado de dominación masculina" (pág. 7).

16 Sobre la idea de "una ciudad sólo existe en las palabras", véase *Republica* 369a.

17 M. Foucault, *The Order of Things: An Archaeology of the Human Sciences*, Nueva York, 1971, pág. 387.

18 Friedrich Nietzsche, *On the Genealogy of Morals*, 1887, segundo ensayo, sec.1; trad. W. Kaufmann y R. J. Hollingdale de W. Kaufmann (comps.), Nueva York, 1967, pág. 57.

19 La frase, en bastardilla en el original, pertenece a J. Elster, *Ulysses and the Sirens: Studies in rationality and irrationality*, Cambridge, 1979, pág. 10.

20 Nietzsche, segundo ensayo, sec. 2; págs. 59-60.

Indice temático

acatamiento, amplio, 241, 301-2
acatamiento, estrecho, 241, 301-3, 306
acatamiento, racionalidad del, 11, 28, 34, 36, cap. 6 passim; y posición negociadora inicial coercitiva, 259, 263-5, 270; y estabilidad, 267; y derechos, 303; y aquiescencia, 306; y elección arquimédica, 350; y sociedad esencialmente justa, 451
acertado y errado, 17, 134
actividad libre (véase también libertad), 126, 129, 130-3, 136, 138, 147; y utilitarismo, 155; y elección arquimédica, 339-41, 344, 349
actor ideal, elección del (véase también persona, el ideal de,) cap. 8 passim
acuerdo en Rousseau, 411
acuerdo racional, 11, 28, 35; y cooperación, 165-6, 169, 179, 198-9; y elección social, 172-3; proceso de, 175-6, 177, 179-80; en Hobbes, 219, 224, 235; y maximización restringida, 235; en Buchanan, 262-5, 308n8; y derechos, 296; y elección arquimédica, 350
acuerdo, hipotético, 27, 29-30, 356, 445
Adeimantus, 354
adhesión ((véase también acatamiento), 29-30
adquisición inicial (véase también apropiación), 11, 270-1, 354-5, 401n15; la opinión de Locke sobre, 271-3
afectividad libre, 456-7, 463
afectos; y valor subjetivo, 76, 78; y locura, 77-8; y valor objetivo, 87, 89; y cooperación, 252-3; y acatamiento, 250; y el hombre económico, 419; y la moral, 427-32
agathos, 408-10, 415, 421
albertanos, 369
alienación, 145, 148
altruismo recíproco, 252-4
amenazas; y negociación, 211-12, 268-70; y racionalidad, 214n21, 250-2
amour propre, 417
anarquía moral, 11; política, 123, 146
animales, 354, 376
anomia, 145
apropiación (véase también adquisición inicial y propiedad privada), 279, 288-90, 366-9, 382-7; y el hombre económico, 41n29, 415, 417-9; de Norteamérica, 387-92
aquiescencia, 306
árabes, 369
argumento de dominación, 233
argumento maximínimo, 121n5
Aristóteles, 132
Arquímedes, 310
Arrow, K. J., 172, 210n6
asociabilidad, definición, 409, 410-1, 416, 419-20, 427
autodesviación, 160-1, 164
autoestima, 350
autonomía, 450-1; y socialización, 349-50, 380, 445, 457-61; e individuo liberal, 454, 455; y restricción, 463
autorización y contribución, en la teoría Rawls, 328-31
Aveling, E., 158n7
Axelrod, R., 255n19
aztecas, 53

Baier, K., 28
Barraclough, G., 41n28
Bell, E. T., 92n28
Benedict, Ruth, 210n5
beneficio individual (véase también beneficio mutuo e interés individual), 30; y maximización restringida, 227
beneficio mutuo (véase también ventaja mutua), 33, 35; y moral, 35, 36, 38; y vínculos sociales, 146; y Hobbes, 162; y la persona justa, 215, 246; v. beneficio individual, 232, 236; y la elección arquimédica, 339-41, 348-9, 350
beneficio relativo maximínimo, 33, 326, 423, 446
beneficio relativo maximínimo, principio del, 33-34, 210-1, 249, 358
beneficio relativo, 209-10
beneficio relativo; véase beneficio relativo maximínimo y beneficio relativo
bien transferible, 206
bien, 47, 80-1; intrínseco, 76-7; y Hobbes, 80, 83, 221-2; como objeto de preferencia, 87, 90; concepción del, 352n29, 448, 454, 455; social e individual, 448
bien, forma del, 39
bienes privados; y mercado, 126-7, 129, 449; y moral, 354, 356, 358
bienes públicos, 442; y moral, 354, 356-9
bienestar material, 380, 390-1, 403
bienestar maximínimo, principio del, 399-400
bienestar social; véase bienestar
bienestar, 148-55, 176-7, 325
Biggins, Brigid, 401n6
Bloom, A., 42n32
Bramhall, Bishop, 432n2
Brandt, R. B., 351n7
bribón sensato, 246
Buchanan, J. M., 261-2, 306, 466n11
Burke, E., 393, 401n16
Burrow, J. W., 465n9
búsqueda y esfuerzo, 434-7, 440, 442; y el individuo liberal, 454-5

cantidad de oportunidades, 366-71, 379-81, 382-7, 388-91, 455
capacidad afectiva para la moral (véase también moral afectiva), 444-5; e individuo liberal, 455; y afectividad libre, 457
carácter acíclico, 68
carácter traslúcido, 236-7; y elección arquimédica, 350
Chamberlain, W., 362-3
ciencias sociales, y racionalidad, 25
clases naturales, perfección de, 82
Cobb, T., 65
coerción, 37, 299, 423; y sociedad, 146-7, 242; y posición negociadora inicial, 259, 263-8; y la división del trabajo, 461
colectivismo, 424
Colón, 384-7
comparaciones de la utilidad interpersonales; véase utilidad, comparación interpersonal.
compensación, 376, 378; de mercado, 283, 287-9
competencia (véase también, mercado perfectamente competitivo), 122, 178, 418
compromiso, 456, 465
comunidad, individualista, 417, 428; e individuo liberal, 465
concesión máxima, 195
concesión minimáxima, 195
concesión relativa minimáxima, 33, 211
concesión relativa minimáxima, principio de la; determinación de la, 198; 33, 35, 187-93, 205, 210, 215-6, 354; argumento para, 195-8; y la maximización restringida, 228-9; y el acatamiento, 241-2; y las amenazas, 251-2; y los derechos, 297-8; y la elección arquimédica, 312, 349; y los bienes públicos, 358-9; y la renta bruta, 361; y la inversión, 397-400
concesión relativa, medida de la, 186-8, 192
concesión, definición, 33, 111-2, 104-9, 192, 194-8; en la negociación de Zethen-Nash-Harsanyi, 199-204
condición admisible de los resultados óptimos, 182, 188, 202
condición natural de la humanidad (véase también estado de naturaleza), 120-1, 125, 160, 162, 217-9, 243, 419-20
conducta racional; véase elección racional
conjunto de estrategias en equilibrio, 107-12
conjuntos de resultados equivalentes, 269
conquista española de los indios, 307
conquista europea de América del Norte, 387-91
consentimiento, y apropiación, 388-9
continuidad, 72-74
contrato (véase también contrato social), 165, 262, 419
contrato social, 311, 350, 417
convención, 41n19, 165
cooperación (véase también interacción cooperativa), 34, 38, cap. 5 passim, 226-7, 305-6, 349, 357, 453; necesidad de, 160-4; racionalidad interna de la, 176, 178-99; racionalidad externa de la, 167, 216, 225-6, 231-42; moralidad de la, 167-8, 204-12; y la disposición a la maximización restringida, 227-31, 240-2; y sacar ventaja, 298-303; social (véase también empresa cooperativa destinada a la ventaja mutua), 361; y orden internacional, 378; intergeneracional, 392-3; y la socialización, 460
cooperacional racional, principio de (véase también concesión relativa minimáxima, principio de la), 11
coordinación, 112
costo de la oferta, 140-1, 359, 362
costo desplazado, 284-6, 303, 385, 392
costos de oportunidad, 359-60, 364-5
creencia, 53-54, 77, 87, 106
Crusoe, Robinson, 131-2, 137, 139, 142, 143, 154, 356, 364
Crusoes, 291-2, 295
culpa, 252-3, 350

Daniels, N., 401n5
Darwall, S. L., 255n20
deber, 17-18, 35, 36, 281-2, 430
debilidad del individuo (véase también insuficiencia del individuo), 404-7
decir la verdad, 212
defección de la estrategia conjunta, 232-3, 237-9
demoktesis, 155
densidad poblacional, 38, 379, 389, 392
derecho de naturaleza, en Hobbes, 217-9, 221-2, 404-5
derechos; y la estipulación, 35, 260, 271-2, 280-3, 287-90, 303; y los recursos disponibles, 141-3, 148-55, 268-9, 280-1, 360-1, 365; y el utilitarismo, 148-55; y Buchanan, 308n8; y la apropiación, 287-90, 368-9, 385-7, 388; y Rawls, 294, 365; y la posición contractual, 295-8; de las naciones, 369-71; de herencia, 394-7
desarme, 119
desigualdad (véase también igualdad), 267; y cooperación, 373-5; e injusticia, 378
desinterés mutuo (véase también no-tuismo); y mercado, 127, 130, 143-8; y circunstancias de la justicia, 161; y la estipulación, 276; y los derechos, 296; y el punto arquimédico, 312, 313; y las relaciones intergeneracionales, 391-5
diferenciación por sexos, 461-2
diferenciación por status, 462-3

digno de confianza, 235
dilema del prisionero; caracterización, 117-8; 11-13, 117-20, 123, 147-8, 164, 231, 404
Dios, 75, 88, 295, 405
disposición a la elección, racionalidad de la, 246-8, 251-2
distribución natural (véase también resultado no cooperativo), 261-5, 267, 306, 371
división del trabajo, 131-2, 289-90, 428, 438-9, 446, 460-1
dobu, 163, 343, 367
dotes básicas, definición, 143, 154, 361; derecho a las, 279-82; y Rawls, 364-5
Dunn, J., 41n26
duración de la vida; véase lapso de vida
Durst, Ruth, 12

Eberstadt, N., 41n28, 381, 401n14
economía, 19-20, 25, 50-51, 122-3, 125-6, 130, 134-5, 223; basada en el acuerdo, 449
egoísmo, 158n6, 223
egoístas, 18, 217, 252-3
ejército de reserva de desocupados, 156
elección (véase también elección racional); en Utopía, 439-40
elección arquimédica, 36, 310-12, 336-51, 446-7, 449-52, 463; y Harsanyi, 316-25; y Rawls, 325-35
elección estratégica, 43, 47; condiciones para, 94, 104, 113-6, 120, 215-6, 227; y mercado, 134; y el Necio de Hobbes, 220; elección de la disposición a la, 231-2, 247-8
elección paramétrica, 43, 125, 232-3, 247-8
elección pública; teóricos de la, 449
elección racional, teoría de la, 19-21, 351; y teoría moral, 20-23; concepciones básicas, 44; y valor, 25; y razón, 48-49, 71-72
elección racional; y preferencia, 44-57; e interés, 58-61; y goce, 61-62; y prudencia, 62-64; condiciones formales de la, 64-74; y valor, 78; e interacción;véase elección estratégica; y teoría económica, 124; y cooperación, 166, 171-80; y maximización restringida, 231-2, 247-8; e imparcialidad, véase elección arquimédica; y moral, 355, 375, 392
elección social, 172-6, 177-8, 183
Elster, J., 40n14, 90n3, 91n12, 91n16, 254, 353n31, 466n19
emotivistas, 18
empeoramiento; caracterización, 213,275; vs. incapacidad de mejorar, 274, 281; e imparcialidad, 290
empresa cooperativa destinada a la ventaja mutua, 28-36, 37-38, 297; y la elección arquimédica, 339-40; y la sociedad esencialmente justa, 447
Engels, F., 159n25, 465n7
equilibrio débil, 106
equilibrio firme, 106
equilibrio no dominado, 107-8
equilibrio reflexivo, 355
equilibrio; definición, 100, 100-3, 107-12; y optimidad, 112-7, 138; y mercado, 122, 129-30, 133, 151, 447; y marxismo, 157-8; anarquista, 264, 306
Er, mito de, 313
escasez; y circunstancias de justicia, 160-1; y adquisición, 272-3; y hombre económico, 422; y Utopía, 434-5; y realización humana, 438-43; e individuo liberal, 454-5
esclavitud del salario, 139
espacio de utilidad, 170, 181, 203
espacio moral, 151, 271-2
espacio-resultado; y cooperación, 169-72; y negociación, 181-2, 185, 194, 203
estabilidad, 153, 242, 267, 306, 420, 452, 456
estado de naturaleza (véase también condición natural de la humanidad), 28, 297, 301, 386-7; y la estipulación, 261, 271, 280, 282, 284, 290; hobbesiano, 271, 279, 281, 384, 411, 434

estado providente, 149, 151
estilos de vida, 354,370,455; criterios de jerarquización, 380-2
estipulación lockeana, establecimiento de la, 276; 11-13, 35, 162-3, 269-308; y apropiación, 287-90, 366-9, 382-7; imparcialidad de la, 290-5; y derechos, 295-8; racionalidad de la, 298-303; y elección arquimédica, 312, 340-3, 349; y renta bruta, 360; y relaciones intergeneracionales, 391, 396-7, 398-400
estipulación lockeana; véase estipulación; loterías lockeanas; y acciones en situación de riesgo o incertidumbre, 69-74, 96-97; y estrategias, 169; y estrategias conjuntas, 174-5; y elección social, 174-5; y elección entre estructuras sociales, 319-20
estrategia conjunta, 169, 172, 176, 179-80, 198, 226-7, 246, 305; y negociación, 181, 206; y maximización restringida, 227-31, 233-4; y moral, 229; y estabilidad, 267; y elección arquimédica, 347
estrategia de amenaza, 269-70
estrategia individual; vs. conjunta, 169, 226; racionalidad de la, 227; y la maximización restringida, 228, 230, 232-3
estrategia mixta, 99; conjunta mixta, 169
estrategia optimizadora justa, 215
estrategia pura, 99; conjunta pura, 169
estrategia; caracterización, 100; y elección racional, 100-4, 116-7; y cooperación, 166-7; y estipulación, 277
Eva, 288-90, 368, 382-7
evaluación (véase también valor), 78
evolución, 252-3
explicación; económica, 19; de la elección y la acción, 54, 87-90; de la conducta altruista, 254; y el hombre económico, 415
explotación, 237-40; y sociabilidad, 29, 460; y marxismo, 155-8; y violación de la estipulación, 378; y moral, 20-21, 354; y hombre económico, 422; y dependencia de los demás, 461
externalidades, 127-8, 135, 164-5, 356; ausencia de las, 127, 138, 141; y negociación, 179; y cooperación, 205, 206-7; y maximización restringida, 250

felicidad (véase también goce), 58
fetichismo, 129
filosofía moral occidental, 408
Foucault, M., 463
Friedl, Ernestine, 466n15
Friedman, M., 358
función de bienestar, 148, 151
función de cooperación, 171-2, 182-3
función de producción, 125, 129, 152, 423

Gauthier, D., 41n19, n21, n29, n30, 112, 121n11, 213n14, 255n18, n26, 352n11, n16, 401n17, n18, 432n6, n16, 433n31
generaciones, relaciones morales entre las, 391-400, 456-7
Génesis, 417
genocidio cultural, 379, 390
Gert, B., 91n21
Gertrude, Queen, 53-4
Gibson, Mary, 412-14, 422
Giges, 403, 405, 430
Glaucón, 27, 38, 354, 403-4, 407-8, 410, 414, 416, 421, 430, 452, 453
Glymour, C., 13
goce, 61-62, 64; y utilitarismo, 82
Goldman, Emma, 159n14
Gompers, S., 415
Gordon, S., 159n11

Gretzky, W., 132, 359-63
Grice, G. R., 28

haberes naturales, en Rawls, 293-5, 328-33
Hare, R. M., 40n13
Harman, G., 86, 91n23, 92n24
Harsanyi, J. C., 21-22, 40n6, n14, 90n4, n5, 91n14, 112, 121n4, n5, n7, n9, 158n4, 176-7, 199, 203-4, 309n10, 312, 316-25
herencia, 345, 355, 395-7
historia de las ideas, 39
Hobbes, el Necio de, véase El Necio de Hobbes
Hobbes, T., 27, 80-83, 121n12, 124, 130, 133, 142, 162-3, 280, 296, 316, 342, 353n31, 405-8, 419-20, 434
Hollingdale, R. J., 466n18
hombre económico, 41n29, 414-32, 440, 453, 465
hombre natural, 408-10
hombre, como criatura y como artefacto, 464
honestidad, 246
Hume, D., 17-18, 30, 43, 77, 90n16, 159n15, 160, 164, 245-6, 314, 316, 343, 405-6, 427, 429
Huxley, A., 321, 424-6

identidad de mercado, 141
identidad personal, 336-7, 449-50
ideología, 39, 306, 413-4
igual racionalidad, 33, 351, 411; y negociación, 196-7, 210-1, 357; y acatamiento, 302-3; y tecnología, 307-8; y justicia, 308; y elección arquimédica, 312-3; y Rawls, 327
igualdad (véase también desigualdad), 290, 357, 448; tecnológica, 308
imparcialidad; y moral, 21-23, 35-36, 82, 351, 416; y negociación, 33, 210-1; y mercado, 137-43, 155; y cooperación, 166-7; y racionalidad, 242, 305, 307; y posición negociadora inicial, 259; y estipulación, 280, 290-4; y elección arquimédica, 310-2, 314, 341, 347; y Harsanyi, 317, 321, 325, 352n14; y Rawls, 326, 329-31, 335
impuesto, 357-8; y renta bruta, 360
incapacitado, 41n27, 354
incertidumbre, 47, 69-74, 96; y elección arquimédica, 316-20, 349
independencia de las alternativas irrelevantes, 203
indios; y la conquista española, 307; y la apropiación europea de América del Norte, 387
individuo (véase también persona); liberación del, 373-5; como producto social, 333-4; aparición y futuro del, 463-4
individuo liberal, caracterización, 453-7; 12, 417, 446, 463-5
instituciones y prácticas sociales, 29-30, 33, 223, 436-7, 458, 460; y elección arquimédica, 310-2, 339-40, 446, 451
insuficiencia del individuo (véase también debilidad del individuo), 426-7, 441-2, 453, 461; y Hume, 405-6
integridad, 66-68, 70-71
interacción cooperativa (véase también cooperación); caracterización, 165-6, 169; como mano visible, 164, 178; representación de la, 169-70; y negociación, 198-9; y concesión relativa minimáxima, 198-9; disposición a la, 215
interacción de mercado (véase también mercado perfectamente competitivo, 32; natural y cooperativa, 160, 164, 179; y renta bruta, 359-60; como interacción asocial, 419
interacción natural, 160, 162, 223, 236, 278; y posición negociadora inicial, 260, 270-1; y estipulación, 280; y transición hacia la interacción de mercado y la interacción cooperativa, 283-90; vs. interacción social, 303
interacción no cooperativa (véase también interacción natural), 166; y negociación, 179; y superávit cooperativo, 183; y persona cooperativa, 215; y derechos, 296

interacción racional; condiciones para, véase elección estratégica, condiciones para; y economía, 122; teoría contractual de la, 261
interacción social (véase también interacción cooperativa, cooperativa y de mercado, y mercado), 146, 260, 333, 334, 420; y posición negociadora inicial, 279; vs. interacción natural, 303; y la elección arquimédica, 326, 336, 342-3, 344; y las relaciones intergeneracionales, 398; y el individuo liberal, 455
interacción, 18; y la teoría de la elección racional, 19-20; y la racionalidad estratégica, 43, 93-5; representación de la, 96-98; y optimidad, 115-6; y elección arquimédica, 339-43
interacción, cooperación y mercado; y la posición negociadora inicial, 258-9, 267, 270-1, 369; y la transición desde la interacción natural, 283-90; y los derechos, 296, 361
interés individual (véase también beneficio individual y preferencia); como término general de las preocupaciones de cada persona, 17-18, 21, 23-26, 33, 122, 464; su diferencia con la preferencia, 45, 56-60, 63-64, 66
internalización de los costos, 286, 303, 385-6, 392
inuit y apropiación europea de América del Norte, 387-90
inversión, 131, 355, 397-400, 457, 464
irracionalidad; y elección, 50, 51, 53-55, 58-59, 63-64, 68; y afectos, 76-77; y condición natural de la humanidad, 120; y creencia, 307; y sociedad competitiva, 413, 422-3

juegos, 435-7, 439, 443
juegos, teoría de los (véase también teoría de la negociación), 11, 12, 20, 25, 33, 121n1, n2
jugadores líberos; definición, 138, 149, 160-1, 206, 212, 271, 277, 294, 389
justicia artificial, 423-5, 446
justicia esencial, 417, 446-9, 451-2, 455-6, 463
justicia instrumental, 446, 452-3
justicia, principios de (véase también beneficio relativomaximínimo, principio de, concesión relativa minimáxima, principio de y principios morales), 21-22, 354; en Rawls, véase principio de diferencia léxica o lexicográfica
justicia; en Rawls, 21-22, 160, 315, 325-335, 364-5; y Glaucón, 27, 403-4, 407; y el beneficio relativo maximínimo, 33; en A. Smith, 122, 124; en Hobbes, 124, 162, 218-9, 222-3, 407-8; y el mercado, 137, 149; disposición a la, 160, 204, 245-6, 333; circunstancias de, 160-64; en Hume, 160-1, 343, 405-7; y la cooperación, 205-6, 308, 453; y la elección arquimédica, 311, 449; en Nozick, 363; y las relaciones de los pueblos, 371, 378, 387, 391; y la razón, 411-3
justicia; y negociación, 33, 205, 210-1; y mercado, 137, 356; y utilitarismo, 151; y maximización restringida, 229-30, 240-2; y recursos disponibles, 271, 293; y aquiescencia, 305-6; y elección arquimédica, 347-8; y sentimientos morales, 350; y participación, 444
justificación y teoría moral, 19-20

Kagan, S., 91n17
Kalai, E. y M. Smorodinsky, 213n14
Kant, I., 23, 143, 146, 314-5, 316, 407, 466n13
Kaufmann, W., 466n18

Lachesis, 313
laissez-faire, 123, 133-5, 147, 149
Lapalace, P. S. de, 88
lapso de vida, 41n28, 379-82, 389, 390
Laver, M., 409
legado, 395-7
lenguaje moral, 17-18, 411, 421
Lévi-Strauss, C., 144

leyes de la naturaleza, en Hobbes, 27, 217-20
libertad (véase también actividad libre); mercado, 32, 130-2, 143, 363-4; natural, 131, 143, 360, 364; en Hobbes, 27, 217-19; en A. Smith, 122, 124; y utilitarismo, 154-5
libertad; racional, 451
líneas de racionalidad, 337-8, 350
Locke, J., 36-37, 260, 271-3, 279, 282, 296, 354, 366, 368
Luce R. D. y H. Raiffa, 73, 90n8, 91n13, 121n1, n10, 213n7, 214n16, n19, 309n9

MacIntyre, A., 12, 408-10
Mackie, J. L., 40n4, 86, 87-89
Macpherson, C. B., 433n20
Malouf, M. W. K., 213n12
mano invisible, 122, 124, 129-30, 290; fracaso de la, 160; y la mano visible, 178; y el pie visible, 222
mantener las promesas, 226
Marx, K., 158n7, 159n23, 353n31, 438-9, 461
marxismo, 141, 155-8, 162-3
maximización (véase también maximimización de la utilidad); y racionalidad, 30, 44-45, 48-49; y moral, 39; y elección racional, 47, 49-52; y explicación de las acciones y las elecciones, 54, 87; del goce, 61; y elección cooperativa, 172-9
maximización colectiva, 148-9
maximización de la utilidad (véase también maximización); y racionalidad, 44, 104-7, 109-10, 246-8, 252, 450; y condiciones para la elección estratégica, 95, 96, 114, 120, 165-7, 215-6; y equilibrio, 100-1; y propiedades de la racionalidad, 113; y optimización, 116-20, 134, 164; y negociación; y las amenazas, 211-2; y el acatamiento de la restricción, 216, 312; en Hobbes, 220-1, 223-4, 226; y la maximización restringida o directa, 227; disposición genética a la, 254; y transferencia improductiva, 266-7; y elección arquimédica, 338
maximización de la utilidad esperada (véase también maximización de la utilidad) y racionalidad, 70-71; y principio de concesión relativa minimáxima, 198; y la estipulación, 277; y la elección arquimédica, 338; y el individuo liberal, 454
maximización del bienestar, 148-9, 155, 176-7
maximización directa; caracterización, 226; 34, 230-1, 243-4, 406; (ir)racionaldiad de la, 231-9; y racionalidad práctica, 247-8
maximización restringida; caracterización, 227-31; 35, 36, 240, 242-5, 375-8, 465; racionalidad de la, 231-40; y altruismo recíproco, 252-4
McDonald, M., 401n15
McGee, Sam, 208-9
mercado perfectamente competitivo (véase también competencia e interacción, mercado); como zona moralmente libre, 31-32, 35, cap. 4 passim; condiciones de la, 124-9, 136; resultado del, 129-30, 133, 141; racionalidad del, 133-6; imparcialidad del, 137-43, 147, 155-8; y utilitarismo, 149-55; y elección arquimédica, 344-5, 349; y hombre económico, 418-9; y sociedad esencialmente justa, 446, 447
mercado, 37, 85, 164, 223, 357, 418-9
mercado, fracaso del, 123, 160, 178
merecimiento, 294
Mill, J. S., 81, 150-2, 320
Miller Brewing Company, 400n3
modelo de equiprobabilidad de los juicios morales de valor, en Harsanyi, 317
monotonicidad, 71-73
Moore, G. E., 87, 91n18
Moore, S., 158n7
moral afectiva (véase también capacidad afectiva para la moral), 407, 428-32, 444-5
moral basada en el acuerdo (véase también posición contractual), 26-28, 35-37, 205-6, 298, 407-8, 428, 464-5; e ideología, 39; y Hobbes, 219; y maximización restringida, 228; y restricción a la interacción natural, 260; y los derechos, 296;

como seudomoral, 414-458; y el hombre económico, 414-7, 453; y el no-tuismo, 431-2; y la sociedad esencialmente justa, 449; y el individuo liberal, 453
moral convencional, 23, 242-3, 354-5
moral de los sentimientos (véase también moral afectiva), 377-8, 392
moral teísta, 316
moral, concepción de la, 19, 23, 39, 313; la moral y la racionalidad, 18-19, 20, 220, 245-6; fundamentación racional contractual para, 27, 28; y el mercado, 32-33, 124-5, 134, 137-43; y el beneficio mutuo, 35, 36-37, 354; y la relación con el valor, 90; y el desinterés mutuo, 143-8, 315-6, 428-31; y el utilitarismo, 148-9; como quimera, 216, 416-17; y el carácter traslúcido, 236, 240; y la satisfacción de los deseos, 249; y la individuación de las personas, 339; como mal necesario, 404-7, 411-417, 421; como instrumento de dominación, 411-4, 417, 422, 431; y la participación, 443-5; valor de la, 452-3
Motor Inmóvil, 265-7, 435
Murningham, J. K., 213n12

Nagel, T., 40n13, 90n10, 255n24, 400n4
Nash, J. F., 103, 107, 180, 203, 214n16, 268
Nash, O., 17
navaja de Ockham, 87
Necio de Hobbes, 206, 217, 219-22, 224-7, 235, 267, 301, 411
negociación, cotidiana, 210-1
negociación, racional, 11, 41n22, 179-80; enfoque general de, 179-94; enfoque formal de, 194-9; condiciones para la, 196; y acatamiento, 216; en Hobbes, 217-8; y los principios morales reales, 229; y el punto arquimédico, 349; y la apropiación, 388; y la inversión, 397-400
negociación, teoría de la, 33-34, 180-1; punto de vista de Zeuthen-Nash-Harsanyi, 180, 199-204, 269-70
Nietzsche, F., 464-5
no nacido, 354
no-tuismo (véase también desinterés mutuo); definición, 127, 409, 423, 441; y fundamento racional para la moral, 415-6, 428-9, 431-2, 444
Nozick, R., 155, 274, 283, 291, 308n1, 309n20, 355, 363-4

obligación, 17, 279; en Hobbes, 217-9
observador ideal, 315
ocupaciones (véase también cantidad de oportunidades), 37
oferta variable, 161-4
opacidad, 236
optimidad; definición, 113; 113-7; y mercado, 122, 129-30, 133-6, 140-1; y externalidades, 128, 298; y utilitarismo, 148-9, 178; y cooperación, 166-7, 171-2; y elección racional, 174-5; y soberanía, 163-4; y maximización restringida, 227-31; y elección arquimédica, 345-6
optimización, 116-7, 120-1
órdenes de preferencia, y elección social, 172-5

pacto, 219-21, 235, 243, 404
parásito; caracterización, 138; 138-9, 149, 160-1, 207, 211, 271, 277
parentesco, 144
Paret, 121n8
Parfit, D., 91n2, 255n20, 256n25, 449, 466n12
participación, 427-8, 429-30, 440-3; y moral, 443-5; y sociedad esencialmente justa, 448; e individuo liberal, 455
pensamiento feminista, 29, 460
pensamiento moral cristiano, 27
pensamiento moral estoico, 27
persona (véase también individuo); concepción de la, 27, 294-5, 338, 422; definición de

la, 126, 136, 141, 262, 290; y posición contractual, 28; los derechos de la, 272, 296; ideal y punto arquimédico, 310-12, 313-16, 324; y utilitarismo, 324-5; y Rawls, 328-35, 338
persona justa, 215-6, 219, 403-4
Platón, 27, 351n3, 400n2, 416, 432n1, n2, n3, 433n29, n32, 464
polis, 407-11, 463
política basada en el acuerdo, 449
política, como restricción, 223-4, 235
Polonia y Solidaridad, 264
posición contractual (véase también una moral basada en el acuerdo), 29-31, 33, 36, 39, 45-48, 423; y Buchanan, 261
posición contractual radical, 415, 419
posición negociadora inicial; definición, 194; 32, 34-35, 180-3, 214n21, cap. 7 passim; en el enfoque se Zeuthen-Nash-Harsanyi, 203, 268-70; en Buchanan, 261-8; y negociación intergeneracional, 398
práctica coordinadora, 29-30
preferencia (véase también interés, individual y utilidad), cap. 2 passim; futura y presente, 62-64, 449-51; y la moral, 143, 145-6; y la negociación, 182-3; y la justicia artificial, 424-5; y la elección arquimédica, 451
preferencia coherente, 46-48, 64-74
preferencia considerada, 45-46, 47-48, 53-64, 66; y valor, 76, 77; e individuo liberal, 454
preferencia expresada, 51-53, 57-59, 61, 116
preferencia manifestada, 51-53, 57, 59, 61, 116
preferencia social, 173-5
preferencia social; véase preferencia
preferencia, medida de intervalos de la (véase también utilidad), 45-46, 47-48, 53-64, 97, 111
preferencia, medida interpersonal de la (véase también comparaciones interpersonales de la utilidad), 148
preferencia, medida ordinal de la (véase también utilidad), 46-47
pretensión, en la negociación; definición, 194, 195-6; la pretensión racional como condición para la negociación, 195; en Zeuthen-Nash-Harsanyi, 199-204; y la negociación intergeneracional, 398-9
principio de diferencia léxica o lexicográfica, 326-8, 330-3, 365
principio de Zeuthen, 112-3, 188-9, 199-2002
principios morales, 19, 20-23, 212, 229, 338
probabilidad; y utilidad, 47, 70-72, 96; y estrategias, 99-100; y maximización restringida, 237-9, 243-4
problema de la negociación, 33
promesa, 226, 464-5
propiedad privada (véase también apropiación); y mercado, 117, 136-7, 155-6
propiedades racionales de los resultados, 42-43, 165
propietario ideal, 316
prudencia, 57, 58, 62-64, 234, 422, 450, 454
psicología moral, 444-5
punto arquimédico, 35-37, cap. 8 passim, 408, 452, 459
punto de amenaza, 269-70
punto de base, 274-5, 301
punto de concesión, 194; condición para la negociación, 195
punto de pretensión, 194

racionalidad de la irracionalidad, 249-52
racionalidad estratégica (véase también elección estratégica), 93-96, 100-1, 104, 106, 130, 164
racionalidad. paramétrica (véase también elección paramétrica), 93, 232
racionalidad, práctica (véase también razón): concepción maximizadora de la, 20, 24-26, 44-45, 65-66, 71-72, 225-6; reinterpretación de la concepción maximizadora

de la, 41n23, 246-8, 252; concepción maximizadora de la y moral, 26, 36; concepción universalista de la, 24-25, 82; instrumental, 48-49, 307, 413, 422-3, 452-3; y la prudencia, 64-64; y la optimización, 116-7; y la negociación, 211-2; la concepción de Hobbes, 221-2; e imparcialidad, 242, 305; punto de vista neokantiano de la, 248; como segunda mejor opción, 250-1; y tecnología, 307; el rechazo de Hume, 406; y la reflexión, 426, 448, 449-50; y la Utopía, 440; y el individuo liberal, 454, 463
Raiffa, H; véase Luce, R. D. y H. Raiffa
Rawls, J., 21-22, 35, 121n5, 127, 155, 160, 164, 243, 293-5, 297, 312, 315, 324, 325-36, 337-9, 351n1, 355, 364-5, 402n19, 440-1
razón (véase también racionalidad, práctica); y moral, 17-18, 24-25, 299, 305-8; e interés, 23; y las pasiones, 43; y los fines de la acción, 48-49; punto de vista de la, 64; y la locura, 77-78; y la cooperación, 160; y Hobbes, 162, 217-8, 220-3; y el punto arquimédico, 310, 314-5; ideal kantiano de la, 314-5, 451; y la justicia, 410-3; y el valor, 452-3; y la naturaleza ideal de la sociedad, 463
razonable vs. racional, 22
razones, prácticas, 24, 226, 248, 465
realización humana, 437-42, 460-1
recursos disponibles individuales, 125, 139, 136-7, 183, 260, 268, 270-1, 272, 279, 294; y la identidad de mercado, 142-3; y los derechos individuales, 151, 155, 279-80, 296, 303; y la herencia, 395-6
redistribución, 35, 37, 300, 360, 364
reflexión; y preferencia, 56, 57-64, 77; y racionalidad, 247-8, 425, 450; e individuo liberal, 454
renta bruta, 140, 345, 354, 359-65
República, 27, 403-4, 433n22, 463
resentimiento, 350, 428
respeto, 350
respuesta maximizadora de la utilidad centroide, 105-6
respuesta maximizadora de la utilidad, 94, 98, 100-1, 103-4, 111-2, 220; y respuesta optimizadora, 113, 115, 120; e interacción no cooperativa, 165-6; y negociación, 180; y la justicia, 215
respuesta optimizadora; racionalidad de la, 115-6, 120; y mercado, 134; y cooperación, 167, 179; y justicia, 215
respuesta racional, 94, 100-1, 104, 105, 110-1, 113-7, 180, 215, 220
restricción, racional y moral, 18, 20, 22, 116, 146-7, 248, 354; justificación universalista de la, 63; fundamento racional contractual de la, 26-27, 34-35; la libertad surgida de la, en el mercado, 32, 122, 133-4, 137-8, 147; impedimiento o rechazo de la, 38, 216, 414-7, 421; y la condición natural de la humanidad, 120; punto de vista kantiano de la, 143; y la cooperación, 179, 199, 205; y Hobbes, 218-9, 223; interna y externa, 223-4; y la maximización restringida, 230-1; a la interacción natural, 260-1, 270-1, 277, 303; dos niveles de, 297-8; y la estipulación, 304, 341-4; y la aquiescencia, 306; como mal necesario, 405-6, 414, 421; y la moral afectiva, 429; aceptación afectiva de la, 431-2; autónoma y heterónoma, 465
resultado cooperativo, 165-6, 172, 192, 215, 228, 236-40, 267
resultado esperado, 100
resultado no cooperativo y posición negociadora inicial (véase también distribución natural), 260, 261, 265, 267, 301, 305-6
riesgo, 47, 69-74, 96; y negociación, 199, 203-4
Rorty, R., 12, 42n33
Roth, A. E., 180, 192
Rousseau, J.-J., 411-5, 416, 417, 420, 427, 461
Rubin, Gayle, 144

sacar ventaja; caracterización, 270-6; y apropiación, 290, 366; y acatamiento amplio, 301-2; y elección arquimédica, 341-2, 347-8; y la apariencia moral, 414
Saltamontes, 434-7, 440

satisfacción de las preferencias generales, medida de la, 318-20, 322-3
satisfacción, 58, 61-62
satisfactorio, 249-51
Schelling, T., 112
Schwier, Ann S., 121n8
Seldes, G., 432n20
Sen, A., 213n8
Shakespeare, W., 90n1
simpatía (véase también simpatizante ideal), 146, 374-5, 377-8, 405
simpatizante ideal, 315-6
sinceridad, 243-4
situación negociadora, 194
Skepticus, 434
Smith, A., 122, 124, 130, 223, 420
Smorodinsky, M.; véase Kalai, E. y Smorodinsky, M.
Sobel, J. H., 13
soberanía del consumidor, 352n14
soberano, 27, 222-3
sociabilidad, 28, 38, 336; y la sociedad como instrumento, 334; y la sociedad esencialmente justa, 461; y la diferenciación por sexos, 462; y el individuo liberal, 464
socialismo, 344-5
socialización, 377, 457; y sociedad de mercado, 142; y capacidad afectiva, 253, 350; y punto de vista instrumental de la sociedad, 334; y elección arquimédica, 310-3, 339-40, 446, 450; y autonomía, 380, 446, 458-9; y justicia artificial, 423-4; y sociedad esencialmente justa, 458-63
sociedad (véase también sociedad esencialmente justa y sociedad de mercado); como empresa cooperativa destinada a la ventaja mutua, 28-30, 32, 296, 339; y anarquía, 123; y desinterés mutuo, 143-8; y estado de naturaleza, 261, 287, 300; e individuo, 294-5, 423-4; y el acatamiento, 301-2; en Rawls, 328-34; y la elección arquimédica, 338-49; intergeneracional, 393, 397-400; como instrumento, 410; en Rousseau, 411, 420; en Hobbes, 419-20; y la afectividad libre, 455-7; y la diferenciación por sexos, 462-3
sociedad de mercado; inversión de, 144-5, 463-4
sociedad esencialmente justa, 417, 445-49, 451, 456-7, 459-63
sociedad homérica, 408
sociedad occidental, 37, 367, 381, 440, 463-4
sociedad utilitarista, 448
sociedades socialistas, 357
Sócrates, 27, 38, 43, 321, 354, 403, 416, 424
sofistas, 27, 46, 408-10, 412-3, 417
Stevenson, C. L., 40n5
Sudáfrica, 264
Suits, B., 434-5
Sumner, L. W., 90n10, 91n11
superávit cooperativo; definición, 181, 185-6, 200, 206-8, 263, 299, 327, 357-8, 366, 373, 399

tecnología; y alcance de la moral contractual, 37; y utilitarismo, 151-2; y racionalidad, 308; y apropiación, 386, 387-91; y Marx, 438; y la justicia esencial, 449; y el individuo liberal, 464
temperatura, 46, 60, 97, 188
teoría de la decisión, 20, 48, 317; bayesiana, 20-21, 70-71, 241
teoría de la utilidad, 74
teoría laboral del valor, 157
teoría moral, 17, 50, 86, 93-94, 116, 143-4; contractual, 11, 26-28, 36-39; y la teoría de la elección racional, 26-28, 36, 350-1; y el punto arquimédico, 310, 350
teoría normativa, 18-20

trabajo, y adquisición, 272, 282-3
transferencia improductiva, 264-6, 299
transitividad, 67-68, 71
transparencia, 235
Trivers, R. L., 252-4
Tucker, A. W., 117

Un mundo feliz, 321, 424-5
unión social, 440-1
utilidad maximínima, 111, 326-28
utilidad, como medida de la preferencia, 45, 46-47, 71, 450; y valor, 45-46, 49, 74, 84; y goce, 61; e interacción, 97; y egoísmo, 110; y mercado, 125; interdependiente, 127; prosecución de la, 133; distribución, 137; y bienestar, 148, 176; y fracaso del mercado, 164; y elección social, 172; comparación interpersonal de la, 97, 148, 176, 177, 185, 191, 318-9, 326; y posición negociadora inicial, 180, 184; y concesión, 185-7
utilidad, esperada, 71-72, 97
utilidad, función de, 71, 94; y mercado, 126-7; y definición de las personas, 126, 136, 142, 422; independiente, 127, 144, 409; y bienestar óptimo, 149
utilidad, maximínima, 111-2
utilidad, promedio, 322-3
utilitarismo, 22, 24, 373; y valor, 82; y mercado, 148-55; y los derechos, 150-6; y la cooperación, 175-9, 295; y la elección arquimédica, 316-25, 327; y la inversión, 399-400
Utopía, 434-7, 439-40, 445

valor absoluto o universal, 48, 79, 82-83; y mercado, 84-130
valor excedente, 156-7
valor instrumental, 436-7, 452-3, 459
valor intrínseco, 75-76, 436-7, 443-4, 454
valor moral, 82
valor objetivo, 48, 59, 76-78, 82-83, 85-90; y mercado, 129
valor relativo, 48, 79-85, 89
valor subjetivo, 48, 59, 75-78, 81-84, 85, 89-90
valor, 74-90; y elección; y preferencia y utilidad, 43-64; y mercado, 84-85, 132-3
valor, concepción teológica del, 75
valor, teoría del, 48, 83; teoría laboral del, 156
valoración asocial, 408-9
Van den Doel, H., 401 n6
vejez, problemas de la, 41n28
velo de ignorancia, 22, 35, 324
ventaja mutua (véase también beneficio y empresa cooperativa para el beneficio mutuo), 28-31; y maximización restringida, 34, 228; y moral, 354; y sociedad occidental, 463-4
Viernes, 139
vínculos afectivos, 145, 147
Von Humboldt, W., 440-1

Wicksteed, P. H., 127, 158 n6, 159 n9
Williams, B., 91 n12
Winch, D. M., 90 n8, n9, 136-7, 149, 151-4, 159 n10, 309 n19

"yo real" kantiano, 338

Zeuthen, F., 112-3, 180, 188-9, 200, 214 n16, 260
zona moralmente libre; y mercado, 32, 35, 123-4, 127, 131, 138, 145; y cooperación, 167

FILOSOFIA

PIER ALDO ROVATTI	*Como la luz tenue*
GEORGES BALANDIER	*El desorden*
HANNAH ARENDT	*Hombres en tiempos de oscuridad*
PAUL RICOUER	*Ideología y utopía*
ERNEST GELLNER	*Cultura, identidad y política*
JEAN FRANÇOIS LYOTARD	*La diferencia*
RONALD DWORKIN	*El imperio de la justicia*
CORNELIUS CASTORIADIS	*Los dominios del hombre*
KARL JASPERS	*La práctica médica en la era tecnológica*
LUDWIG WITTGENSTEIN	*Sobre la certeza*
JEAN FRANÇOIS LYOTARD	*El entusiasmo*
JEAN FRANÇOIS LYOTARD	*La posmodernidad (explicada a los niños)*
TZVETAN TODOROV	*Frágil felicidad*
GIANNI VATTIMO	*El fin de la modernidad*
GIANNI VATTIMO	*Introducción a Heidegger*
JACOB BRONOWSKI	*Los orígenes del conocimiento y la imaginación*
NORBERTO BOBBIO	*El problema de la guerra y las vías de la paz*
GILLES DELUZE	*Empirismo y subjetividad*